The Dissertation Committee for Sung-Kwan Yang
Certifies that this is the approved version of the following dissertation:

Teachers' Perception of Use of Student Performance Information: Technology Acceptance Model

Committee:

Jay D. Scribner, Supervisor

Charles T. Clark

Lori J. Einfalt

Pedro Reyes

Laura M. Stapleton

Teachers' Perception of Use of Student Performance Information: Technology Acceptance Model

Sung-Kwan Yang, B. A.; M. A.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

December 2003

Dedication

To my loving wife, Seung-Hyun Son,

and my lovely sons, Sunoo Yang & Jinoo(?) Yang

To our parents

Woo-Seok Yang & Hak-Pyo Kim

Pyung-Up Son & Jung-Sook Kim.

Acknowledgements

"Why have I found such favor in your eyes that you notice me - a foreigner?"

(Ruth 2:10)

I give God all the praise for supporting me through this experience. My accomplishments and blessings are a reflection of His holy plan.

I am extremely appreciative of the support I received throughout this process from my supervisor, Dr. Jay Scribner, and director of the Educational Productivity Council, Dr. Terry Clark. Dr. Scribner has been a true source of encouragement since my time in the Educational Policy and Planning Program. I am grateful for both his guidance and his friendship. Dr. Clark has been unfailingly gracious and generous with his time and humor giving me valuable feedback during each successive step of the research process.

My thanks to Dr. Lori Einfalt for standing by my side as a friend, colleague, counselor, and committee member for five years. My appreciation to Dr. Laura Stapleton for her enthusiasm for this study and assistance in finalizing this work. My sincere gratitude to Dr. Pedro Reyes for his mentorship in research methods during the early phase of graduate school.

Thanks also to my colleagues at the Educational Productivity Council for their assistance over time with this project. In particular, I want to thank Brendan Maxcy for his great critical mind and thoughtful input, Shang-Min Choi for his kindness in assisting me with formatting this dissertation, Joe Dinan for his editorial support, and Shin-Hyun Ka for his technical support.

Many people have contributed to the completion of this project by through their love and encouragement. I would like to thank our parents and all of our family who have always given me love and support. The church family at the Korean Baptist Church of Austin including Senior Pastor Sung-Bae Kim has all been spiritually supportive and patient during this endeavor and I am grateful to them.

My sincere appreciation goes to each member of my thesis committee: Dr. Hyung-Kwan Kim for serving as an excellent advisor during my master program, Dr. Ki-Hang Wang for his interest and continued encouragement for my study, and Dr. Ki-Sung Ahn for his guidance and vision of my study.

Finally, my warmest appreciation goes to my wife, Seung-Hyun, for her helping me realize a dream through unconditional love, patience, and many hours of loneliness. This project would not have come true without her undying support and understanding. My son, Sunoo learned that there is a “doctor” not working at hospitals. I hope he will soon know the meaning of the word.

Teachers' Perception of Use of Student Performance Information: Technology Acceptance Model

Publication No._____

Sung-Kwan Yang, Ph. D.

The University of Texas at Austin, 2003

Supervisor: Jay D. Scribner

The purpose of this study was to examine (1) what type of student performance information is available to teachers; (2) how they use the information; and (3) what factors impact positively and/or negatively on teachers' use of student performance information. This study also examined (4) the relationships among perception of usefulness, ease of use, attitude, intention to use student performance information, and external pressure to use the information.

This study employed both survey research and individual focused interview methods. The perceptions of teachers' using student performance information were investigated through a survey of 206 teachers from 16 campuses

in an urban school district in Texas. Data analysis included structural equation modeling

Major findings from research question one were: The school district with well-developed assessment systems responded in an organized way to rising standards by improving alignment within their local curriculum. Schools and classrooms also were engaged in ongoing initiatives to improve student achievement under their own circumstances.

Major findings from research question two were: Teachers used the benchmark assessment information (1) to check the efficacy of local curriculum and instructional practices; (2) to assess state curriculum standards and/or objective mastery, and (3) to prepare for state mandated tests. (4) Teachers reported they didn't use the benchmark assessment reports, or used at the minimum level as directed by the school district.

Major findings from research question three were: Teachers reported identification of student needs, alignment of curriculum and tests, preparation for instruction, information access, and information quality positively impacted their use of benchmark assessment information. Teachers reported poor quality of test and reports, user-unfriendly format, low information access, time consumption, external pressure, and misalignment with curriculum negatively impacted on their use of the benchmark assessment information.

Major findings from research question four reported that the teachers' perception of intention to use student performance information was affected directly and/or indirectly by perceived usefulness, perceived ease of use, attitude, perceived information quality, and external pressure. Only the user-guide rating variable in individual difference category showed a statistically significant relationship with perceived usefulness, intention to use, and perceived information quality.

Table of Contents

List of Tables

Table of Figures

CHAPTER ONE

INTRODUCTION

BACKGROUND

The effort to restructure education in schools across the United States has generated many proposals for reforming the entire education system. Prominent among them, standards-based school reform has been considered to be one of the major vehicles for facilitating change. In the year 2000, forty-nine states had standards in place, forty-four of which established standards in English, mathematics, social studies, and science. Thirty-eight states have revised or developed new standards during the last two years. Twenty-one states have developed assessments intended to test whether students meet standards in all four of the above-mentioned academic subjects (Fuhrman, 2001).

One of the most basic methods for achieving standards-based accountability is public reporting of student performance. The public is able to use this information to determine the effectiveness of their schools and as a basis for making demands regarding improvements. Most states and school districts are providing performance reports to their communities that include student academic achievement and other measures, such as graduation rates, attendance, and post-secondary outcomes (Goertz, 2001).

The underlying assumption for reporting data on student performance is that such data will be helpful to educators, parents, and others who have an investment in public schools. Information about achievement from formal test assessments, in particular, is thought to provide important data for academic decision making. This enables all of the stakeholders to do their part toward improving instruction and student learning (Linn, 2001). Teachers and administrators are urged to analyze and use student performance data to directly inform school decisions about how curriculum and instruction can be changed to yield better learning outcomes for low-performing students (Goertz, 2000). Parents and other interested individuals use student performance results to exert pressure on local educational systems by demanding more from low-performing schools.

The Use of Student Performance Information

While providing performance data to schools and communities has gained centrality in policy rhetoric, has it had any discernible effects on schooling practices? Massell (2001) responds to this question by examining the ways in which states and districts promote the use of data and whether and how these data affect educators' attitudes and practices. State and local education agencies promote the use of data focusing on school improvement planning procedures that

encourage data-based decisions. Efforts are made to design internet systems that make data accessible and user-friendly, and to provide professional development to improve skills in data analysis and application. Thus, Massell (2001) showed how state and local efforts such as these affect local educators in two primary areas: (1) creating positive attitude toward the utility of data as a script for guiding practice, and (2) increasing skills in the application of data in local decision making.

Other studies describe consequences associated with student test results, such as generating greater pressure on teachers and administrators and instituting questionable behavioral changes among teachers, and narrow test preparation practices (Cannell, 1987; Linn, Graue, & Sanders, 1989; Romber, Zarinnia, & Williams, 1989; Shepard, 1990). In addition, Marso and Pigge (1999) provided several examples relating to questionable effectiveness of public dissemination of test information. In some schools, educational administrators tend to avoid conveying standardized test results to their teachers. Even when test results are conveyed to teachers, the results typically are unavailable until six or eight weeks after the administration of the tests. Consequently, classroom teachers often report a low value and very limited use of the results in their day-to-day classroom instruction (Marso & Pigge, 1999).

Massell (2001) provided the reasons for teachers' perception of little value for student performance data.

> First, the majority of classroom teachers are simply receiving direction from others about how to proceed, or are expected to tease out the implications of data for teaching on their own. Second, teachers are likely to think that data gathering and analysis tasks interfere with the 'real' work of schooling. Third, the assessments generated by states and supplemented by districts, schools, teachers, and others often create an overwhelming amount of information to which teachers and administrators are expected to respond. (p.165)

Thus, the assumption that reporting student performance is helpful to the teachers who play the most basic role in improving instruction and student learning raises a few salient questions. For example, how much information does the report of student assessment contain the method used to report student performance test information improve its usefulness to teachers? How do certain teacher characteristics influence the use of student performance data in making instructional decisions for students (e.g., understanding and internalizing how a particular accountability system works; acquiring the knowledge and skills concerning data interpretation; and motivating the individual teacher to use test results to make instructional decisions)? Finally, what impact do organizational factors have on the use of student performance information by teachers (e.g., external demands to raise test scores by shifting to instructional planning around test content; and internal pressures to create new/different professional norms)?

Applying a Conceptual Framework

A core set of theoretical frameworks emanate from a body of research on innovation adoption and information technology acceptance. These frameworks would appear to be heuristically relevant to seeking answers to the preceding questions because the effective use of student performance information represents a significant innovation for public educators. The adoption of a systematic approach to data-driven instructional improvement efforts is at the heart standards-based reform.

Diffusion of Innovations (Rogers, 1983), the Theory of Reasoned Action (Ajzen & Fishibein, 1980), the Technology Acceptance Model (Davis, Bagozzi, & Warshaw, 1989), the Theory of Planned Behavior (Ajzen, 1985: Taylor & Todd, 1995), and Social Cognitive Theory (Compeau & Higgins, 1995) have received widespread validation for many technological innovations. For the purposes of this research investigation, data-driven decision-making based on the use of student performance information is dependent upon the technological innovation of information management.

Among the above mentioned theories, Rogers' (1983) Diffusion of Innovation (DOI) theory and Davis' (1989) Technology Acceptance Model (TAM) have received the most attention in the Information System (IS) literature on individual-level innovation adoption frameworks (Gallivan, 2001).

The Diffusion of Innovations (Rogers, 1983) and Technology Acceptance Model (Davis, et al. 1989) identify *perceived attributes of an innovation* as key predictors explaining adoption, and feature *user intention to adopt a technology* as the dependent variable. Rogers (1983) identifies five perceived attributes of innovation as influencing adoption behavior: 1) relative advantage, 2) complexity, 3) compatibility, 4) trialability, and 5) obervability. *Relative advantage* captures the extent to which a potential adopter views the innovation as offering an advantage over previous ways of performing the same task. *Complexity* refers to the degree to which an innovation is perceived as relatively difficult to understand and use. *Compatibility* encompasses the degree to which an innovation is perceived as being consistent with the existing values, needs, and past experiences of potential adopters. *Trialability* measures the extent to which potential adopters perceive that they have an opportunity to experiment with the innovation prior to committing to its usage. Finally, o*bservability* indicates the degree to which the results of an innovation are visible to others.

Davis (1989), on the other hand, explains adoption behavior by suggesting only two perceived attributes that influence innovative adoption: *perceived usefulness* and *ease of use*. *Perceived usefulness* is defined as the extent to which a person believes that using a particular technology, in this case student performance data, will enhance job performance, while *perceived ease of use* is

defined as the degree to which a person believes that using a technology will be free from effort. Davis' concept of *perceived usefulness* is analogous to Rogers' concept of *relative advantage*. Likewise, Davis' concept of *perceived ease of use* is similar to Rogers' reference to *complexity*. Clearly, Davis' Technology Acceptance Model offers a more focused framework for predicting technology acceptance and use by potential adopters.

As TAM (Davis, 1989) offers two attributes compared to the five attributes set forth in Roger's (1983) Diffusion of Innovations framework, the researchers (Agarwal & Prasad, 1997; Moore & Benbasat, 1991; Tornatzky & Klein, 1982) found that a smaller set of beliefs could predict current usage behavior, as well as future use intentions. Therefore, a more parsimonious set of beliefs have keen selected for this research. TAM, the more parsimonious model, will be employed as the main model combining elements from other frameworks to explore factors impacting *teachers' perceptions of intention to use student performance information*.

It should be mentioned, however, some researchers assert that these frameworks neglect the realities of implementing technology innovations within organizations, especially when adoption decisions are made at the organizational level, rather than at the individual level (Fichman & Kemerer, 1997; Orlikowski, 1993; Wynekoop, 1992). In this kind of non-voluntary adoption process, the

reality of innovation adoption and implementation within organizational settings will require modifications to the framework, as proposed in the present study. Thus, an attempt will be made to apply insights from a growing body of research literature associated with innovation adoption and information technology acceptance to the contemporary phenomena where teachers are called upon to use student performance information in a mandated environment to make instructional decisions for students. Hence, an adaptation of the Technology Acceptance Model that accounts for organizationally mandated adoption decisions can serve as a valuable tool for exploring those factors that affect teacher intention to use student performance information to make instructional decisions.

CONCEPTUAL MODEL

This study presents and tests a model (see Figure 1) of teachers' intention to use student performance information in a suburban school district in Texas. The research model suggests that the teachers' perception of intention to use student performance information is affected directly and/or indirectly by: (1) individual difference; (2) perceived information quality; (3) perceived usefulness; (4) perceived ease of use; (5) attitude; and (6) external pressure.

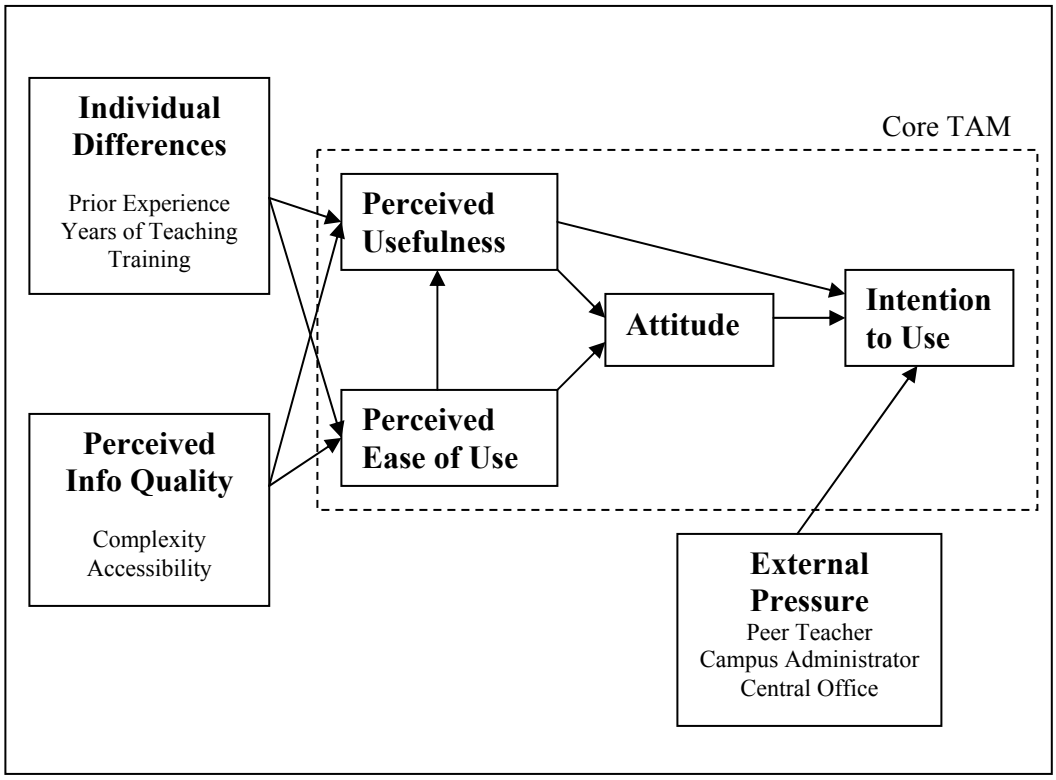


Figure 1: The Conceptual Model

Individual Level of Analysis

The Technology Acceptance Model (TAM) has been extensively applied and utilized in the studies of innovation adoption and diffusion at the individual level (Agarwal, 1999; Davis, 1989; Gefen, 2000; Lucas & Spilter, 1999; Venkatesh & Davis, 2000). TAM, adapted from the Theory of Reasoned Action (Fishbein & Ajzen, 1975), posits that use acceptance is determined by perceived usefulness and perceived ease of use. The Theory of Reasoned Action (TRA) articulated by Fishbein and Ajzen (1975) explains people's actions by identifying

the causal connections between various components: beliefs, attitudes, intentions, and behavior. TRA is concerned with the determinants of actual and intended behavior. Davis (1989) adapted the TRA model to develop the TAM by replacing the attitudinal determinants of TRA with two distinct variables-perceived usefulness and perceived ease of use. Perceived usefulness is defined as the extent to which a person believes that using a particular technology (performance information) will enhance job performance (teaching and learning), while perceived ease of use is defined as the degree to which a person believes that using a technology will be free from effort. This study attempts to investigate the impacts of two basic components in TAM.

TAM is a theorem of belief formation process which provides a basis for understanding how individual differences might influence the development of teacher beliefs about the use of technology. McGoch & Irion (1952) propose a law of proactive inhibition or interference that individuals' prior knowledge and experience interfere with their ability to learn to exhibit specific behaviors. Zmud (1979) derived a taxonomy of individual difference variables, which included the categories of demographics, personality, and cognitive style. Other research has examined individual difference variables such as user involvement (Jackson et al., 1997), training (Compeau & Higgins, 1995; Venkatesh & Davis, 1996), and prior experience (Thompson et al., 1994). This study will examine the impact of similar

prior experiences and knowledge (Gick & Holyoak, 1987; McGeoch & Irion, 1952), the length of tenure in the workplace (Harrison & Rainer, 1992; Nickel & Pinto, 1986), and participation in training (Igbaria, Gamers, & Davis, 1995) on the perceived usefulness and ease of use of student performance information.

The perception of usefulness and ease of use of a technology can be influenced by the quality of technology in itself as well. Extended TAM (Venkatesh & Davis, 2000) posits that the perception of output quality is expected to explain significant unique variance in perceived usefulness. This study regards complexity and accessibility as the quality of information. The complexity of the testing information is defined as the level of sophistication of the information system in terms of the number of variables provided. Accessibility is defined as the level of acquisition of information by teachers to use the student performance data to set goals, measure progress, and discuss strategies.

Organizational Level of Analysis

It has been theorized in technology adoption and diffusion models that social norm (i.e., culture or climate of emphasizing accountability, performance management, and statewide-standardized test system) are an important determinant in the intention to use the technology. Social norms are defined as a "person's perception that most people who are important to him think he should

or should not perform the behavior in question" (Fishbein & Ajzen, 1975, p.302). From the organization's perspective, how to foster an environment such that norms could be utilized to influence one's intention to use certain information becomes a critical issue when facing the challenges of information adoption and diffusion.

Obviously, one way to induce technology usage is to simply mandate the use of the information technology. Agawal and Prasad (1997) suggested (tense) that while mandating technology use can provide the impetus to overcome inertia associated with a new behavior; such efforts may not be sustainable over time. Moreover, given some evidence that mandating technology use against the explicit will of an individual may result in negative consequences (e.g., Ram & Jung, 1991), it would be desirable for managers to promote voluntary acceptance of information technologies. To distinguish between mandatory and voluntary use of student performance information, this study posits external pressure to use the information as a moderating variable.

STATEMENT OF THE PROBLEM

Accountability and standards-based reform controversies suggest that schools will improve if educators have better information about academic goals and outcomes of students that they serve. Richards (1988) emphasized "the

regular collection of information, evaluation of that information, and institutional action that resulted from that evaluation (cited as in Fleck, 1999, p. 16)." Elmore and Rothman (1999) stated that the basic standards-based reform model will include a concept of school improvement by instructional change aligning with 'student performance data'. Public reporting student performance data is the most basic form of accountability and standards-based reform (Goertz, 2001).

According to Massell (2001), the roles of promoting data by the state government and school districts have changed into more intense focus on school data and new assessment measures. School districts, at the same time, have promoted more and more the use of student performance data through professional development, school improvement planning, and additional incentives for data use.

However, in the real world, there are many barriers to the process of promoting the use of student performance data. For example, teachers are less likely to use the performance information to improve instructional practice if they don't have enough time to involve in the school improvement planning process (Ogawa & Collum, 2000; Porter & Smithson, 2001). Teacher's attention to testing in instructional planning and delivery tends to be dissipated due to lack of capacity-building conditions, such as feedback on student assessment measures

and results, meaningful professional development, and structured teacher collaboration (Goertz, 2001).

What seems to be problematic, even in a well designed model, is that this kind of intensification of data use always includes the pressure of accountability system toward educators. Those barriers seem to be closely tied to educators, especially teachers, who are frustrated by various regulations, replacing staff or reducing school autonomy as a sanction, so to speak, educators '*being held* accountable'.

PURPOSE OF THE STUDY

This research explores the relevance of the Technology Acceptance Model by analyzing individual teachers' perceptions of use of student performance information, particularly as it is also displayed in the format of various standardized test reports. This model will be relevant to describing facilitators and barriers to promoting teachers' use of student performance information.

This study examines the limitations inherent in applying the individual perceptions of use of the student test reports in a school in which there is a school district mandate to adopt these reports. After modifying the core Technology Acceptance Model (TAM), which is mainly focused on the teacher's perception of usefulness, ease of use, attitude, and intention to use of student performance

information, this research will extend the core TAM by adding the external pressure construct which may explain some of the individual teachers' intentions to use the test reports within a mandatory adoption situation.

RESEARCH QUESTIONS AND HYPOTHESES

This study will be guided by four research questions. The first three questions will pertain to type of student assessment, actual use of student assessment information, and school culture regarding assessment.

1. What types of student performance information are available and accessible to teachers?
2. How do teachers use the testing information?
3. What factors influence teachers' use of student performance information?

This study will examine the relationships among perception of usefulness, ease of use, attitude, intention to use student performance information, and external pressure to use the information.

4. Is the Technology Acceptance Model a useful conceptual tool for understanding the relationships among variables associated with teachers' intention to use student performance information?

The following hypotheses are related to the fourth research question.

- H1: Perceived usefulness and ease of use will mediate the influence of selected individual difference and perceived information quality variables on attitude and intention to use student performance information.
- H2: Attitude will mediate the influence of perceived usefulness and ease of use on intention to use student performance information.
- H3: Perceived ease of use will have a direct effect on attitude and an indirect effect on attitude through perceived usefulness.
- H4: Perceived usefulness will have a direct effect on intention to use and an indirect effect on intention through attitude.
- H5: External pressure will have a direct effect on intention to use student performance information.

DEFINITION OF TERMS

Data, Information, and Knowledge. *Data* are objective facts, presented without any judgment or context. Data becomes *information* when it is categorized, analyzed, summarized, and placed in context. Information develops into *knowledge* when it is used to make comparisons, assess consequences, establish connections, and engage in a dialogue. Knowledge can be seen as information that comes laden with experience, judgment, intuition, and values. (Empson, 1999)

Information Technology. *Information technology* refers to the collection of products and services than turn data into useful, meaningful, accessible

information. This study will examine student performance information as a means for considering student instructional needs. This study will see the student performance information as a type of information technology.

Perceived Ease of Use (PEU). Davis et al. (1989, p.985) define perceived ease of use as "the degree to which the prospective users expect the target system to be free of effort." This concept will check whether it is easy for teachers to learn to interpret the student performance information, and to interact with other faculty members in a clear and understandable way.

Perceived Usefulness (PU). Davis et al. (1989, p.985) define perceived usefulness as "the prospective user's subjective probability that using a specific application system will increase his or her job performance within an organizational context." In this definition, PU is linked to whether teachers will ultimately gain rewards from their use of a report including student performance information and enhance the attitudes they have toward using the report.

Perceived Information Quality (PIQ). PIQ refers to complexity and accessibility as the quality of information. The complexity of the testing information is defined as the level of sophistication of the information system in terms of its number of variables provided. Accessibility is defined as the level of acquisition of information by teachers to use the student performance data to set goals, measure progress, and discuss strategies.

Beliefs and Attitudes. *Beliefs* are an individual's cognitive evaluation of the consequences of a particular behavior (i.e., the use or acceptance of an information technology artifact) while *attitudes* reflect an affective response to the behavior in question (i.e., a generalized liking or disliking for the behavior).

Intention to Use (IU). *Intention to use* is the willingness of teachers to use the student performance information during work activities.

External Pressure. *External Pressure* reflects the impacts of three interrelated social forces, peer teachers, campus administrators, and central office staff, impinging on an individual facing the opportunity to adopt or reject a new system. This construct is similar to subjective norm, voluntariness, and other social influence factors.

Standards-Based Reform. According to U.S. Deputy Secretary of Education Marshall Smith, standards-based reform is typically based on state-level reforms that implement more rigorous content and performance standards across grades and discipline. It requires that curricular material and assessments be aligned with these standards. Preservice teacher education and teacher professional development must also support these goals.

SIGNIFICANCE OF THE STUDY

This research will expand on previous research addressing theories on information technology adoption. The results of this study may expand theoretical development and a practical perspective in the following ways:

a) This study may posit and find empirical support for a theory of how mandated use of student performance information drives teachers' intention to use performance reports including student information and its influence on attitudes about the new reports.

b) This study may identify factors that can be instrumental in facilitating the use of student reports through their positive influence on usefulness and ease of use. Campus and district level administrators might consider the usefulness and ease of use of student performance information driven by standard-based reform. When teachers perceive the use of this information to be organizationally mandated, school administrators should take into account the need for training and the timely dissemination of student performance information.

c) Educational administrators must encourage educators to consider the critical role played by the *management of knowledge* in efforts to support educational reform, specifically current standard-based reform. Educators who work on campuses, in districts, or within state education agencies face mounting

pressure to demonstrate the measurable effects of their educational practices to the public at large.

This pressure can be reduced by incorporating information technology as a component of a knowledge management system in the implementation of high-stakes accountability across the nation.

LIMITATIONS

This study is not a search for cause and effect on 'behavior'. It focuses on exploring factors that affect the 'intention' to use student performance information as a predictor of actual usage. As indicated in the first limitation, actual usage of student performance information is not measured because data are gathered at single point in time and not longitudinally. The model in this study utilizes intentions as a dependent variable. Like Diffusion of Innovation, future research postulates that many different outcomes are of interest in information technology adoption, including the initial decision to use student performance information and the continued or sustained use of that information.

This study will employ cross-sectional research designs. Although this study is operationally easier to conduct, it doesn't support truly causal inferences. In future research, longitudinal studies should be conducted to closely examine the causal structure of the overall frame espoused in Figure 1 on page 9.

CHAPTER TWO

REVIEW OF THE LITERATURE

INTRODUCTION

As this study involves technology acceptance in the standard-based reform movement, it is important to review and frame this research contribution relative to prior studies. The review of the literature will examine three areas of research: 1) consideration of current trends and features of accountability and standards-based school reform movement; 2) exploration of general research streams, such as, Information Diffusion Model, Theory of Planned Behavior, Technology Acceptance Model (TAM), and Information Technology (IT) adoption; 3) discussion of TAM in more depth, including previous studies and constructs used for technology adoption studies, consistent with the proposed model for this study of modified TAM.

STUDENT PERFORMANCE INFORMATION UTILIZATION

Call for Accountability

Concept and Framework on Accountability

Researchers offer definitions of accountability that differ in slight but meaningful ways, even though they have limitations. Browder (1975, as cited in Kirst, 1990) concluded that there are no commonly agreed on definitions; as a concept, accountability needs refinement; and accountability has become highly politicized.

People seem to become accountable if they are held responsible for something and, as a consequence, are "subject to the obligation to justify that something" (Popham, 2000, p.283). There are two major focal points of Popham's definition-the object of the responsibility and the justification process of the responsibility. Rothman (1995) defined educational accountability as "the process[es] by which school districts and states attempt to ensure that schools and school systems meet their goals" (p.189). His definition also has similar focal points: processes and goals. A focus on processes raises questions about the accountability tools or mechanisms. A focus on goals indicates that the intention of accountability processes is to influence schools and districts toward the accomplishment of their goals, and their success in doing that is the most important criterion on which to judge their value. Macpherson (1996) also argued that criteria and process in education are essential components of the accountability system. The criteria are used to make judgments about

performance and services, whereas process means to collect, store, report, and use data to improve the quality of performances and services.

Wagner's conception of accountability (1989) is framed as a response to five questions. (1) What level of accountability is to be provided? The answer would require not only description and explanation but also justification. (2) Who is expected to provide the account? The answer would relate to identifying the person or parties responsible for the act creating it, such as, teacher, principal, school district, etc. (3) To whom is the account owed? Parents and students typically get to a formal account of teacher performance from a written report of student achievement. (4) What is to be accounted for? The answer would be closely linked to the goals (Rothman, 1995) and the criteria (Macpherson, 1996) of the accountability system. For instance, student achievement, collaborative professional cultures, high expectations for student achievement, and clear goals are all accountability measures (Creemers & Reetzig, 1996). (5) What are the consequences of providing an account? The answer would prescribe the characteristics of the system, such as voluntary vs. obligatory, rewards vs. punishment, or support vs. pressure.

The last four of Wagner's questions are substantially the same as the four "attributes" that Adams and Kirst (1999) used to distinguish among different types of accountability in their classification scheme.

> Types of accountability are distinguished by four attributes. First, accountability systems express different relationships between principals and agents. Principals are those who establish an expectation (regarding a task to be accomplished) and to whom an account is owed; agents are those of whom performance is expected (in accomplishing the task). The second, third, and fourth attributes indicate the nature of the accountability relationship, or what is expected of agents, the type of mechanism employed to ensure accountability, and the nature of the incentive used to compel agent's actions." (Adams & Kirst, 1999, p.467)

Based on these attributes, Adams and Kirst identified six types of accountability: bureaucratic, legal, professional, political, moral, and market accountability. Other researchers have similar forms of accountability systems. For example, Darling-Hammond (1988) provided five similar types of accountability, with the exception of moral accountability. Kogan (1986, as cited in House, 1993) identified three models of educational accountability; state or public control, professional control, and consumer control. Garn (2001, p.578) completed an extensive review of the types of accountability systems above-mentioned:

1) Bureaucratic accountability is based on procedural compliance with established standards and regulations evaluated by local, state, or federal bureaucrats analyzing compliance reports and/or monitoring at the school site (Cuban, 1988; Darling-Hammond, 1988; Kirst, 1990).

2) Performance accountability is based on data from various indicators that may be used to stimulate action, monitor compliance, and include rewards or sanctions (Levin, 1974; Rivera, 1994).

3) Market accountability is based on student/parental choice among schools. … Market accountability is measured by consumer participation and could be recognized through waiting lists or attendance records that indicate increases or decreases in student enrollment (Chubb & Moe, 1988, 1990; Darling-Hammond, 1988; Kirst, 1990).

4) Professional accountability is based on the demonstration of educators to their peers that they have the appropriate knowledge, values, and skills to ensure competence and serve the public interest (Darling-Hammond, 1988; Firestone & Bader, 1992; Rivera, 1994).

History of Accountability

Although educational accountability issues recently gained nationwide attention as a tool of educational reform, they have a long history. For example, "in the 1817 Georgia law applying to 'poor schools'… the commissioners were forbidden to pay a teacher any salary if an examination showed that his students had not made good progress in that quarter (Georgia Education Law, 1965, as cited in Lessinger, 1970, p.114)." The U.S. Department of Education was formed in 1867 to collect and report educational statistics (Ogawa & Collum, 2000). Spring (2000) argued that the accountability movement was an attempt to restore power to professional educators, who have been threatened by the community control movement and the public criticism of schools in the 1950s.

Lessinger (1970) considers the community control movement a threat to the quality of education. Proposing an analogy from medical practice to oppose the concept of democratic control, he applies it to education emphasizing

professional knowledge gained through research and study. In return, he recognizes the responsiveness of the public by reporting the results of schooling, which is the essence of the accountability movement. His principle of public accountability includes the adoption of educational engineering in order to answer to the public in terms of results, students' competence, and monetary investment. This notion of *accounting* for competence and cash is grounded on "the simple long-established notion that the public has a right to know what benefits they are receiving for their tax dollars (p.31)."

When Lessinger (1970) proposed this kind of accountability, the schools had already spent billions of dollars on giving disadvantaged students financial assistance to help them study with students from richer families, based on Title I of the Elementary and Secondary Education Act (ESEA) in 1965. Therefore, before the 1980s, accountability was grounded in performance measurements, incentives, results linked to resources, and capacity-building. In short, accountability through better management and fiscal procedures (Lessinger, 1970; Pipho, 1989) was alive and well. Pipho (1989) displayed that the main topics of accountability during the 1970s were assessment of student achievement, evaluation of programs, setting goals for education, specifying objectives for learners, PPBS (planning, programming, budgeting systems), MBO (management

by objectives), MIS (management information systems), uniform accounting systems, and performance accreditation systems.

Beginning in the 1980s, citizens demanded the academic performance of America's schools in keeping with the management and financial accountability strategies of the 1970s (Adams & Kirst, 1999; Kirst, 1990). *A Nation at Risk: The Imperative for Educational Reform* (National Commission on Excellence in Education, 1983) attributed the success of America's schools, that is, their capability to promote the nation's prosperity, security, and civility, to their academic performance. The commission recommended curriculum alignment, high expectations, rigorous and measurable standards, time on task, and teaching quality.

From the mid-1980s through the 1990s, new demands for educational accountability symbolized the nation's commitment to educational quality. For example, President Bush's America 2000 strategy defined national educational goals and called for national and state report cards to track student achievement. President Clinton's Goals 2000: Educate America Act promoted content and performance standards and student assessments to measure progress toward those standards (Adams & Kirst, 1999). No Child Left Behind Act of 2001 focused on the major provisions of performance standards, adequate yearly progress targets, reliability issues about individual school results, and some validity issues on

information system to assess and improve educational quality (Linn, Baker, & Betebenner (2002).

In short, the movement of accountability went though restoring power to professional control from community control: The movement also concentrated on values such as desegregation, compensation for perceived socioeconomic disadvantages, inclusion of handicapped and limited-English-proficient populations, resource equity, and service delivery in the 1960s. During the 1970s, growing performance measurements, incentives, results linked to resources and capacity-building through better management and fiscal procedures were touted. Since educational accountability shifted from *process* to *outcomes* in the 1980s (Elmore, Ablemann, & Fuhrman, 1996), the movement continuously shifted from districts to schools, from compliance regarding inputs and practices to student performance, and from comparative performance to performance against a standard of achievement (Elmore, 1999; Fuhrman, 1994; Hansen, 1993; Murnane & Levy, 1996; O Reilly, 1996).

Trends in Student Assessment

The conceptual and historical review of educational accountability suggests the following major accountability components: defining goals or standards, allocating authority, managing incentives, building capacity, measuring

progress, reporting results, and enforcing consequences, all of which are related to student performance. In the 2000s, these components bring a conclusion of standards-based reform. Standards-based reforms sound like school accountability which consists of "high expectations for performance; expectations regardless of students' backgrounds; alignment of assessments to the standards; permission to make the instructional and structural changes; pre and in-service professional development; and accountability to create incentives and assistances to schools (Elmore, 1999, Fuhrman, 2001)."

Past components of accountability systems and current standards-based reform are associated with student testing, standard setting, and monitoring by reporting the results of student performance.

The Testing Movement

The testing movement includes standard-based reform and monitoring issues. In the early 20th century two test-related developments occurred that critically impacted the use and development of standardized tests. The first was the IQ test, designed by Afred Binet for French school children. The second event was the advent of a specific testing method, the multiple-choice item, invented in 1914 by Frederick J. Kelly. Tests were given and scored quickly and impartially

by untrained personnel. (Clarke, Madaus, Horn, & Ramos, 2000;Haney, 1984; Spring, 2000). According to Lauren Resnick and Daniel Resnick,

> Standardized tests in various school subjects were introduced into American schools in the period 1880-1920 when booming enrollments, large school-building programs, and the cult of efficiency in industry combined to encourage the schools to justify their performance in quantitative ways to local taxpayers. Short-answer and multiple-choice tests were viewed as cost-efficient and objective measures in which there might be some public confidence. (pp.13-20)

In 1925, the College Board chose a panel of experts to guide development of a multiple-choice test for determining admittance into college. As a result, the first Scholastic Aptitude Test (SAT) was administered to 8,000 candidates in June 1926 (Haney, 1984). The invention of the high-speed scanner in 1955 heralded a new age in testing. The combination of the high-speed scanner and the multiple-choice format produced accurate and cost effective methods for testing a large population (Clarke et al., 2000). From the 1950s to the present day, international competitors impacted on increasing desire of standardized testing for accountability movement. For example, there has been the general discontent with the quality of education that began when the Soviet Union launched Sputnik in 1950s, followed by the basic skills movement in the 1970s. Japan and Germany demonstrated higher productivity in the economy of the 1980s, and was followed by excellence movement in education to the present. By the mid 1980s, California was incorporating content standards into curriculum frameworks and textbooks

(Fuhrman, 2000), which called for tests covering core subjects such as mathematics, English, and science. Efforts to link assessment to such standards were promoted by professional publishing companies like the *California Test Bureau, PsychCorp, ACT Riverside Press, and Scantron* (Clarke et al., 2000).

Throughout the 1990s, standard-based reform employed two major types of tests (Popham, 2000). Comprehensive nationally standardized achievement tests are distributed by commercial testing companies, such as *California Achievement Tests, Comprehensive Tests of Basic Skills, Iowa Tests of Basic Skills, Metropolitan Achievement Tests, and Stanford Achievement Tests.* These tests are norm-referenced tests to be able to compare a student score with the national norm group. Popham (2000) assumed that each state selected one of the tests which fit well with the current curriculum. The second type of tests is an exam designed by state to check the attainment of minimum performance which is called a 'criterion-referenced test. Reporting often includes breakdown of performance data by student groups according to ethnicity and socioeconomic status.

> A good many of the tests that form these state-customized assessments are now described as "standard-based" tests because they are ostensibly aligned with a state's approved content-standards, that is, the body of knowledge and skills that state's educational policymakers hope will be taught to the state's students. (p. 286)

Standards-based Reform Controversy

Throughout the 1980s and 1990s, while standards-based reform has been in vogue, the effectiveness and potential problems of the reform have been also hotly debated. The following questions might be placed at the center of the debates. Does this movement have an impact on teaching and learning, thus resulting in increasing student performance? Is it an effective device for enhancing equal opportunity for underprivileged groups of students? For some researchers, standards-based reform offers an important means of improving student achievement, as well as giving more educational opportunities to minority students. However, others such as Porter (1994), are concerned about the standards-based tests.

> Those who… fear that standards will trivialize education and de-skill teaching by being too prescriptive… will create an inflexible delivery system that will be incapable of coping with differences between poor schools and rich schools, able students and weak students, well-prepared teachers and teachers teaching out-of-subject. (p. 430)

There are always double-edged features in various educational policies. In the arena of accountability, there are some contentions within the standards-based reform debate, the high-stakes test debate, the state-mandated test debate, etc. The heart of debate pertains to the test in itself. Since the testing movement received the spotlight, the question has been raised about the validity and reliability of standardized tests. Concerns exist regarding the validity of the test, narrowness of

content, neglect of higher-order thinking skills, and the limited relevance and meaningfulness of multiple-choice format (Baker, 1989; Herman, 1989; Shepard, 1990). Generally speaking, there is no perfect test administration. A test score is a product that combines what a test purports to measure and error of measurement and external error such as student's health condition on the day of the test. Based on these concerns, some researchers have raised an issue about whether increasing test scores actually means significant improvement of students' learning (Cannell, 1987; Linn, Graue, & Sanders, 1990; Shepard, 1990).

The second debate is about the impact of the limited features of standardized test on teacher's behavior and belief system. Haney (2000) and other researchers argued that standardized testing resulted in narrowing of the curriculum and instruction, fostering of anxiety, confusion, fear, shame, anger, and/or mistrust, deskilling of teachers and a perception of powerlessness, and loss of instructional time due to test preparation and testing (Brown, 1993; Romberg, Zarinnia, & Williams 1989; Smith, Edelsky, Draper, Rottenberg, & Cherland, 1989; Smith, 1991).

Another concern examines the relationship between standardized test and equity in education. Some studies identify common themes that standards may heighten educational inequities. Students attending low performing schools are at a disadvantage and have the least access to the instruction needed to reach high

standards (Orfield, 1999). Teachers in urban and poor school systems are the least likely to receive the necessary professional development resources (Alington & Walmsley, 1995). Students who are socio-economically disadvantaged may have to repeat a grade or may drop out of school because of the intensified testing system (Gordon and Reese, 1997; Haney, 2000; Hoffman, Pennington, Assaf, & Paris, 1999; Kozol, 1991; McLaughlin, Shepard, & O'Day, 1995). After all, without considering more fundamental educational needs and adequate educational resources, such as the staff or the materials to support a curriculum based on advanced content, there are still unresolved problems in the era of educational quality, outcomes, and excellence. The process and resources issues, which were the centerpiece of the 1970s' debate on educational accountability, suggest that standards-based reform is destined to confront heated disagreement above mentioned.

The debate on equity issues aligning with the *use of test*, such as systemically reporting students' performance to the public, has been more controversial in the state of Texas. Some have suggested that under the system, Texas reduced performance gaps more rapidly than other states (Grissmer & Flanagan, 1998) and that the system played a crucial role in promoting equity by: 1) focusing district leadership and public attention on performance disparities, and 2) holding schools and districts accountable for the performance of all students

(Skrla, Scheurich, & Johnson, 2001). Others suggested that the quick gains reported by the state were misleading (Klein, Hamilton, McCaffrey, & Stecher, 2000), the high failure rates visited on students of color were discriminatory (Haney, 2000; Natriello & Pallas, 1998), and that the system would ultimately harm the students and schools it purports to help (McNeil & Valenzuela, 2000).

There is a question about whether the utilization of test results may influence schools as a support mechanism or pressure device. Reporting the results of student assessment to the public who acquired it at the cost of delivering their control over education to the professional educators (Lessinger, 1970) has been historically a major component in the educational accountability system and standards-base reform movement. On the one hand, the reporting system assumes that test data can be used as feedback to shape classroom instruction, and that it ultimately contributes to continuous improvement of schools and school districts (Glatthorn, 1994; Rozum-Pratto, Gontarz, Flint, & Thomas, 1997). On the other hand, some researchers view the reporting system not as a supportive tool, but as a bureaucratic prescription for instruction (Darling-Hammond, 1997; McDonnell & Elmore, 1987; Porter, Smithson, & Osthoff, 1994). Recent studies have provided an integrated interpretation. Barber and Phillips (2000) reported the 'fusion' of pressure and support criticizing the polarizing argument as useless.

Wilson and Floden (2000) also presented a mixed picture of teachers who vary in their interpretation of standards reform and in how they change their practice.

In summary, the controversy about standards-base reform movements exhibited problems of validity and reliability of measures, positive and negative impact of test oriented policy on teacher's behavioral and psychological aspects, limited capacity for establishing social justice, and the utilization of student assessment information relating to resource issues and structural issues. These five debates are epitomized by three fundamental issues, 1) lack of alignment, 2) equity, and 3) capacity (Goertz, 2001). The first issue, lack of alignment, considers how well standards and assessment have been aligned. Regardless of test types, such as state-customized test or norm-referenced test, there is generally a time lag between standards and tests. With regard to the equity issue, standards-based reform is intended to contribute to improve achievement for all students. Valid performance data should be established for 'fair' comparison of all groups of students. The final issue is about capacity-building conditions. As Fullan (2001) states, the standards-based reform focused primarily on the adoption level by intensifying pressure and support, leaving a question mark on implementation. It is necessary to dig into capacity-building conditions, such as timely feedback of student assessment, the leadership of campus and district administrators, and

professional development in order to implement standards-based reform at the campus level.

EFFECTIVE USE OF STUDENT PERFORMANCE INFORMATION

The above-mentioned call for accountability and standards-based reform controversies suggest that schools will improve if educators have better information about academic goals and outcomes of students that they serve. Richards (1988) emphasized "the regular collection of information, evaluation of that information, and institutional action that resulted from that evaluation (cited as in Fleck, 1999, p. 16)." Massell (2001) also considered "the *outcome data* as an integral part of the improvement process (p.148)" in the state accountability system. Elmore and Rothman (1999) state that the basic standards-based reform model will include a concept of school improvement by instructional change aligning with 'student performance data'. After all, public reporting of student performance data is the most basic form of accountability and standards-based reform (Goertz, 2001). Furthermore, Goertz (2001) submits;

> Public can then use this information to demand improvements in their schools, or possibly to choose alternate schools for their children. While nearly all states report annually on student performance, about a dozen states use public reporting as their primary accountability mechanism. In most of these states, districts must administer and report the results of a statewide assessment. (p.44)

But while reporting student performance data has been placed on the conceptual and historical center of accountability, what impact on schooling practices, such as teacher's instructional decision making and teaching practice and school improvement planning, has accountability generated? This section begins to answer this question by reviewing how student performance information has been used by administrators in the state government, school districts, and at the campus level and what the major implications for teachers are.

Limited Use of Student Performance Information

Generally speaking, tests are used to provide feedback to individual students (formative) or to make decisions about grades, promotion, and graduation (summative). To expand these notions of use to the campus level, test results sometimes are used to diagnose and prescribe treatments for emergent problems (Guthrie, 1990; Oakes, 1989; Smith, 1988). They are used for evaluating the effectiveness of schooling practices (Nuttall, 1994), which ultimately serves to inform the public about the improvement of student learning.

Historically, the accountability and standards-based reform have shown the latter: summative orientation aligning with *pressure* to educators by public side. This is the reason that many states have tried to hold schools or districts accountable through 'performance monitoring' (Burstein et al., 1992; David;

1987; Nuttall, 1994; Selden, 1994), a mechanism for providing rewards or sanctions similar to "market-like" situation (Richards, 1988). This function has accompanied the use of student performance information for evaluating the effectiveness of polices or programs (Nuttall, 1994). Actually, this function of program evaluation has been required since establishment of the ESEA of 1965, where school districts were expected to report yearly the test results of Title I students (Heubert & Hauser, 1999).

At the theoretical level, the use of a bureaucratic control mechanism, such as performance monitoring and program evaluation for rewards and sanctions, still embraces the standards-based school reform. According to Massell (2001), the roles of promoting data by the state government and school districts have changed into more intense focus on school data and new assessment measures with procedural mechanisms like school improvement planning supported by data-driven decisions, user friendly data reporting by Internet, and professional development. School districts, at the same time, have promoted more and more the use of student performance data through professional development, school improvement planning, and additional incentives for data use. This approach continuously reinforces the theory that measurement, report of test scores and other indicators by internet, and rewards and punishment will inspire schools to

help all students attain the minimum level of state standards as depicted in diagram.

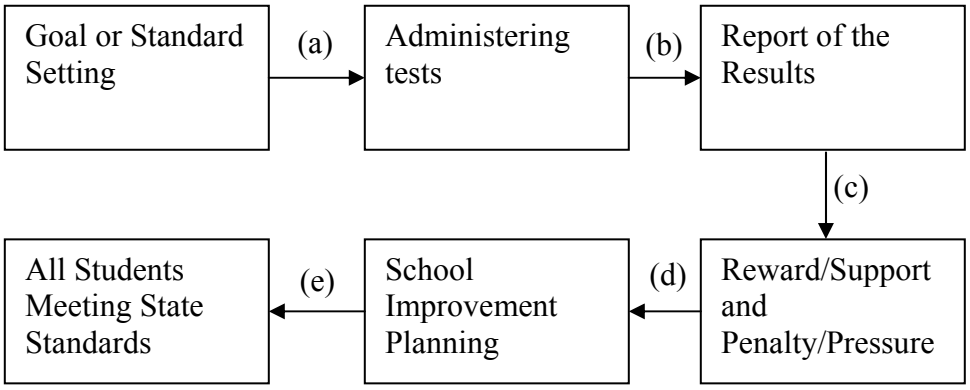


Figure 2: Six Common Steps and Five Connection of Standards-based School Reform

However, in the real world, there are many barriers to the process of linking those six entities. For example, (a) concerns about how well states have aligned their standards and assessments (Goertz, 2000) and (b) the validity and reliability issues in the testing administration as described in the previous section, and (c) policy and public expectations for increasing student achievement have emerged faster than the technical capacity leading to public misperception of the state assessment practices (Stiggins, 1995). Under this kind of situation, administrators are likely to use the test scores for the purposes in a slightly different way. Olson (1998) took an example of *Iowa Test of Basic Skills* test, which is nationally norm-referenced, and used it to identify low-performing schools and students. (d) Some studies critique that administrators and policy

makers do not reliably and effectively use data to make decisions and adopt policies (David, 1988), and site-based councils typically do not carefully analyze or consciously utilize needs assessment data to develop school improvement plans (Malen, Ogawa, & Kranz, 1990). Furthermore, teachers are less likely to use the performance information to improve instructional practice if they don't have enough time to be involved in the school improvement planning process (Ogawa & Collum, 2000; Porter & Smithson, 2001). (e) Teacher's attention to testing in instructional planning and delivery tends to be dissipated due to lack of capacity-building conditions, such as feedback on student assessment measures and results, meaningful professional development, and structured teacher collaboration (Goertz, 2001).

What seems to be problematic, even in a well designed model, is that this kind of intensification of data use always includes the pressure of accountability system toward educators. Those barriers, in every step of standards-based school reform efforts, seem to be closely tied to educators, especially teachers frustrated by various regulations, replacing staff, or reducing school autonomy as a sanction, so to speak, educators '*being held* accountable'.

New Approach to Account-Able

What makes it possible to change from the approach of educators 'held accountable' with the inherent sanctions of traditional educational governance into an approach of educators being "account-able?" Some researchers have focused on site-based management with shared governance and *teachers' leadership* instead of centralized decision-making with hierarchical leadership (Elmore, Peterson, & McCarthy, 1996). Others have considered a *civic-minded* approach which is based on locally developed public-private partnership instead of 'policy minded' policies and regulations mandated at the federal, state, and district level (Bracey, 1999; Fiz & Gorard, 2000). Also, the *inquiry-minded school,* which incorporates issues on standards and assessment into its culture, is touted to have the ability to generate accountability from within the school not outside (Rallis & MacMullen, 2000). In a similar vein, Baker and Richards (1999, as cited in Fleck, 1999) offered an ecological perspective on educational management in which knowledge is constructed, learning is reciprocal and collaborative, and the constant auditing of indictors in the internal and external environment allows the system to anticipate the future.

After all, these studies seem to be overlapped in two key components: 1) teachers and school administrators' capacity-building and 2) shared

accountability, leading to educators being able to account for their schooling. Those components are melted into the meaning of being "account-able."

> To be accountable means to be obligated to understand and explain one's actions. Accountability relies on feedback; it links performance with results. Thus accountability in schools is not only about *results* but also about every aspect of *teachers' actions*. … Put simply, practitioners who are accountable evaluate their own practice and then use the information to improve. (Rallis & MacMullen, 2000, p. 769)

This meaning of educators being accountable places focus on interpreting student's performance information and using it to improve schools rather than on simply gathering data and reporting it. It also seeks to increase teachers' engagement through "reflective inquiry to define the educational goals by setting their *own* standards or by taking *ownership* of external standards (Rallis & MacMullen, 2000, p. 770)," in other words, internal accountability.

In order to achieve internal accountability, student performance information must comprise certain qualities. According to LeMahieu (2000), the information must be understandable and meaningful to all participants. In fact, this quality relating to validity and reliability issues should precede any other efforts to make it possible to success in school changes. The second quality of information for achieving internal accountability is about emphasizing context, which means that information should include a description of the context and conditions, suggest causes, and describe educational processes, not just a single

outcome variable, such as test scores. Third, information must be enriched and not simplified or reduced to indicators or statistical composites. Last, information must be available through *information infrastructure*. Therefore, those qualities require the development and operation of sophisticated information management systems for monitoring outcomes, explaining goal achievement, and identifying useful interventions (Scheerens, 1991).

According to Ogawa and Collum (2000), information management systems would require high technical and financial costs for dealing with the enriched indicators. There are a lot of data relating to schooling: demographics on students, courses in which they enroll, programs to which students are assigned, attendance figures, dropout rates, parent responses, standardized test scores, results of teacher-made tests, results of performance assessment, and so on. Ogawa and Collum (2000) described these data in terms of 'density,' the number of variables in the report, and 'complexity,' the number of relationships between those variables that the educational system examines. Whereas performance monitoring tends to use relatively less dense and complex set of indicators for holding schools accountable with high feasibility, the information management system can afford to operate with much more complex and dense set of educational indicators, and to be employed by all the participants including teachers and administrators, which would have a limited feasibility. Many states

adopt least dense-highly feasible performance monitoring systems as a strategy for improving the academic performance of schools (Selden, 1994). This adoption is supported by the argument that more extensive systems would be unmanageable and overly complex (Blan, 1993; Shavelson et al., 1989), and high cost (Oakes, 1986).

Even though there exist those advantages in the performance monitoring system, if the information from large-scale, standardized tests isn't really meaningful to teachers, if it is infrequent and not diagnostic, if it measures competency on the standards but not the prerequisites needed to attain competency, or if it is not individualized for students at different levels, then it will not be perceived to be meaningful by educators. Popham (2000) provided a couple of teachers' responses for the lack of meaningfulness. .

> Some teachers simply pay no attention to any curricular content, irrespective of its intrinsic worth, if that content is not directly assessed by the state's accountability tests. Other teachers snare a copy of the test, then teach their students to master the test's actual items or remarkable similar items. Still other teachers have been caught giving out copies of an "easy-to-follow" correct-answer key as students are about to do battle with a high-stakes accountability tests. (p. 287)

It seems that such information is more useful for administrators than for teachers. That is why there is some research on principal's use of data but not on teachers' use of the information. Schein (1992) emphasized the role of leadership in handling information, such as importing information efficiently, moving it to

the right place, making internal transformation, getting feedback, and repeating the whole cycle of information collection, as necessary for an organization to deal effectively with turbulent environments. Fleck (1999) explored how elementary school principals use assessment information in order to determine what constitutes effective practice. Furthermore, the principal traditionally viewed as a problem-solver and decision-maker (Leithwood & Steinbach, 1995), and professional evaluator (Glasman, 1994; Stufflebeam, 1996, as cited in Fleck, 1999) must be competent in collecting and using evaluation results for decision-making and leadership tasks.

However, the campus is a microcosm of educational system. This study will explore the most critical of factors, teachers' use of student performance information based on the belief that "teachers are at the center of all efforts to improve schools, that without their full participation and leadership, any move to reform education-no matter how well intended or ambitious-is doomed to failure (Lieberman & Miller, 1999, p.20)."

Research on the relationship between teachers and the use of test information overall tends to rely heavily on teachers' perceptions about the impact of the test movement (Brown, 1993; Glasnapp, Poggio, & Miller, 1991; Grant, 2000), not about utilizing the information constructively. For instance, some studies indicate that the effects of statewide testing are different between

high stakes and low stakes (Heubert & Hauser, 1999; Madaus, 1988; Smith, 1991). Other studies explore the impact of teachers' subject matter knowledge, teachers' views of learners and teaching, and grade level on the teachers' belief and practice relating to mandated state tests (Glasnapp, Poggio, & Miller, 1991; Grant, 2000; Smith, 1991). District expectation, local context and condition, and state and district policy climate also emerge as significant factors to the teachers' perception on the test movement (Brown, 1993; Firestone, Mayrowetz, & Fairman, 1998; Grant, 2001). Furthermore, some studies focus on organizational structures such as aligning curriculum with teaching, providing appropriate training, and testing to increase accountability for student progress (English, 2000; Glatthorn, 1994; Rozum-Pratto, Contarz, Flint, & Thomas, 1997; Olson, 2000).

As the vast majority of test information employed in state-level accountability systems do not give teachers adequate targets for effective instructional planning (Popham, 2000), most studies above mentioned typically turn to the impact of external variables on teachers' perception and attitude about test information. Therefore, there remains a question about how teachers can use such test data and other indicators effectively in their teaching.

In summary, the major practical challenge of current accountability and standards-based school reform is to identify and expand the role of educators in

an active and positive way, such as an approach that ensures educators are "*account-able*", rather than merely "*held accountable*." This approach may help teachers and principals engage in collaborative inquiry and action for the success of school improvement. As a result of their commitment to the process, account-able educators will engage in formative evaluation, assessment-curriculum alignment, and politically sophisticated campus improvement planning for students' learning.

UNDERSTANDING INDIVIDUAL ACCEPTANCE

A core set of theoretical frameworks emanate from a body of research on innovation adoption and information technology acceptance. These frameworks would appear to be heuristically relevant to seeking answers to the preceding questions because the effective use of student performance information represents a significant innovation for public educators. The adoption of a systematic approach to data-driven instructional improvement efforts is at the heart standards-based reform.

The issue of individual acceptance of information technology has been researched from multiple theoretical perspectives using a wide range of constructs and definitions. Figure 3 presents a broad synthesis of existing theorizing about this phenomenon.

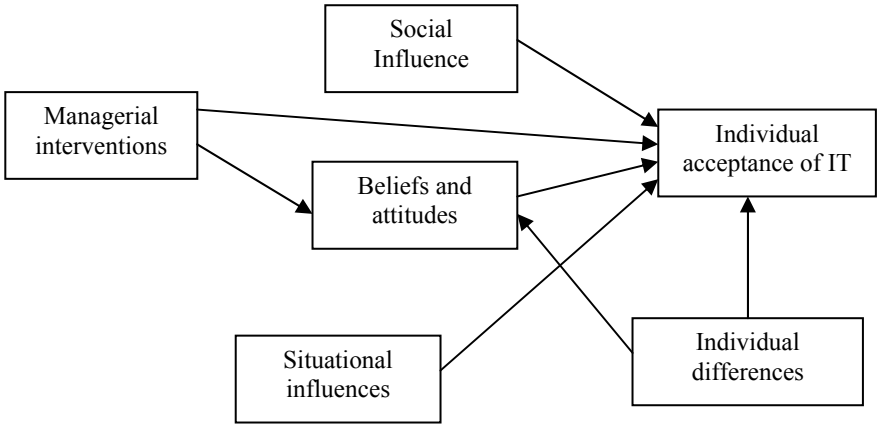


Figure 3: The Phenomenon of Individual Acceptance (Adapted from Agarwal, 2000, p. 86.)

According to Agarwal (2000), a majority of the conceptualizations of IT acceptance have drawn on robust theories from social psychology, notably the theory of reasoned action (TRA), the theory of planned behavior (TPB), diffusion of innovations (DOI) theory, and social cognitive theory (SCT). TRA and TPB were formulated as generalized explanations of a broad range of individual behaviors, including the use of information technology. Both theories posit that such behavior is influence by an individual's intention to perform the behavior. Intention is determined by the individual attitude toward the behavior and the subjective norm. Subjective norms capture an individual's assessment of the extent that referent others would desire the performance or nonperformance of the focal behavior, while attitude captures a general affective response toward the behavior. Attitude in both theories is influenced by an expectancy formulation of

beliefs about the consequences of performing the behavior weighted by an individual's evaluation of each consequence.

TRA has been empirically tested and has found support in the context of the acceptance of information technologies (Taylor & Todd, 1995; Davis, Bagozzi, & Warshaw, 1989). Drawing upon the theoretical foundations of TRA, Davis (1989) proposed the technology acceptance model (TAM). In the TAM as in the TRA, attitudes predict intentions, and intentions predict behavior. Unlike TRA, TAM does not include a subjective norm component as a determinant of intentions because of its uncertain theoretical and psychometric status (Davis et al., 1989).

Rogers' (1983, 1995) theory of the diffusion of innovation has informed a considerable body of IT acceptance research. This theory, constructed through a meta-analysis of a variety of innovations in divers contexts, may be characterized a rich and complex information-centric view of innovation acceptance. The essence of this theory suggests that innovation adoption is a process of uncertainty reduction. Individuals gather and synthesize information about a new IT from the social system within which they are situated. This information processing results in the formations of beliefs about using the IT. Beliefs cause individuals to accept or reject the artifact; that is, beliefs are the drivers of the decision to adopt.

Empirical tests of Rogers' conceptualization in the domain of IT have largely supported the major prediction of the theory (Bracheau & Wetherbe, 1990).

A final theoretical frame-social cognitive theory (SCT)-has been utilized recently by IT researchers not specifically to predict acceptance behaviors but to provide additional insights into the determinants of acceptance behaviors. Social cognitive theory is an outcome of over two decades of research by Bandura and his colleagues (Bandura, 1997). The essence of SCT rests in the notion of triadic reciprocity, in which individual behavior is posited to be an outcome of a complex set of interactions between individual characteristics and environmental and situational factors: Behaviors, individual differences, and situational contingencies mutually influence and affect each other. Compeau and Higgins (1995) posited and found support for positive effects of computer self-efficacy on computer usage, affect, and outcome expectations related to performance, a construct conceptually equivalent to perceived usefulness.

It is important to note that all of these theories explicitly or implicitly apply to behaviors that are under the volitional control of individuals. However, there are situational impediments. Generally speaking, one way to induce technology usage is to simply mandate the use of the IT. Agarwal and Prasad (1997) suggested that while mandating technology use can provide the impetus to overcome initial inertia associated with a new behavior, such effects may not be

sustainable over time. Moreover, given some evidence that mandating technology use against the explicit will of an individual may result in negative consequences (Ram & Jung, 1991), it would be desirable for managers to promote voluntary acceptance of information technologies.

TECHNOLOGY ACCEPTANCE MODEL

Beliefs and Attitudes

For conceptual clarity, Figure 3 encapsulates beliefs and attitudes as a single construct. Beliefs are an individual's cognitive evaluation of the consequences of a particular behavior (i.e., the use or acceptance of an information technology artifact) while attitudes reflect an affective response to the behavior in question (i.e., a generalized liking or disliking for the behavior). In the disaggregated view, attitudes are composed of beliefs, affect, and conation. Affect connotes emotional engagement with the concept; and conation represents the action taken in response to beliefs and affect. Each three components is an integral aspect of attitudes. The one-dimensional view conceptualizes beliefs, attitudes, and behaviors as three separate constructs that are causally related. In both perspectives, the three components are posited to be consistent with each other: Positive beliefs about a concept such as using an IT are likely to coexist with positive affect and subsequent manifestation of the corresponding behavior

of using the IT. A majority of the research in IT acceptance has adopted the one-dimensional view of attitudes, treating beliefs, attitudes, and behaviors as distinct and separable construct.

Perceived Usefulness/ Perceived Ease of Use

The Technology Acceptance Model (TAM) has been extensively applied and utilized in the studies of innovation adoption and diffusion at the individual level (Agarwal & Prasad, 1999; Davis, 1989; Gefen, 2000; Lucas & Spilter, 1999; Venkatesh & Davis, 2000). TAM, adapted from the Theory of Reasoned Action (Fishbein & Ajzen, 1975), posits that use acceptance is determined by perceived usefulness and perceived ease of use. The Theory of Reasoned Action (TRA) articulated by Fishbein and Ajzen (1975) explains people's actions by identifying the causal connections between various components: beliefs, attitudes, intentions, and behavior. TRA is concerned with the determinants of actual and intended behavior. Davis (1989) adapted the TRA model to develop the TAM by replacing the attitudinal determinants of TRA with two distinct variables-perceived usefulness and perceived ease of use. Perceived usefulness is defined as the extent to which a person believes that using a particular technology (performance information) will enhance job performance (teaching and learning), while perceived ease of use is defined as the degree to which a person believes that

using a technology will be free from effort. This study attempts to test the impacts of two basic components in TAM

External Pressure

It has been theorized in technology adoption and diffusion models that social norm (i.e., culture or climate of emphasizing accountability, performance management, and statewide-standardized test system) are an important determinant in the intention to use the technology. Social norms are defined as a "person's perception that most people who are important to him think he should or should not perform the behavior in questions" (Fishbein & Ajzen, 1975, p.302). From the organization's perspective, how to foster an environment such that norms could be utilized to influence one's intention to use certain information becomes a critical issue when facing the challenges of information adoption and diffusion.

Obviously, one way to induce technology usage is to simply mandate the use of the IT. Agawal and Prasad (1997) suggested that while mandating technology use can provide the impetus to overcome inertia associated with a new behavior, such efforts may not be sustainable over time. Moreover, given some evidence that mandating technology use against the explicit will of an individual may result in negative consequences, it would be desirable for managers to

promote voluntary acceptance of information technologies. To distinguish between mandatory and voluntary use of student performance information, this study posits social pressure to use the information as a moderating variable.

Individual Differences

The term individual differences can be interpreted most generally to connote dissimilarities among people including differences in perceptions and behaviors, traits and personality characteristics, and variables that connote differences attributable to circumstances such as education and experience.

Reviewing and synthesizing the research literature on IT implementation and use, Zmud (1979) notes that individual differences can be categorized into three classes: (1) cognitive style, the mode of functioning shown by an individual in his or her perceptual and thinking behavior; (2) personality, the cognitive and affective structures maintained by individuals to facilitate adjustments to events, people, and situations encountered; and (3) demographic/situational variables, such as sex, age, experience, education, and professional orientation (p.95). Studies have shown that individual differences (e.g. gender, age, education, and professional orientation) play an important role in the how information technology is used (Zmud, 1979). In an analysis of diffusion research, Rogers (1995) found that early adopters of an innovation had higher socioeconomic status

than later adopters. Status was typically indicated by such variables as income, education and occupational prestige.

Among various individual differences, the constructs of 'self-efficacy' and 'personal innovativeness' have been identified with strong theoretical underpinnings. Self-efficacy refers to individual's beliefs about their ability and motivation to perform specific tasks (Bandura, 1977, 1986). In the domain of information technology, studies of the effects of self-efficacy collectively point to its crucial role in determining individual behavior toward and performance using information technologies (Compeau & Higgins, 1995, Gist, Schwoerer, & Rosen, 1989). According to Agarwal and Prasad (1998), personal innovativeness in the domain of information technology is defined as "the willingness of an individual to try out any new information technology (p. 206)." Empirical results suggest that personal innovativeness moderates the relationship between beliefs and intentions. Other research has examined individual difference variables such as user involvement (Jackson et al., 1997), training (Compeau & Higgins, 1995; Venkatesh & Davis, 1996), and prior experience (Thompson et al., 1994).

Davis et al. (1989) suggested that the internal psychological variables (i.e., the beliefs) that are central to TAM fully meditate the effects that all other variables in the external environment may have on an individual's use of an innovation. They noted, for instance, that "External variables… provide the

bridge between the internal beliefs, attitudes, and intentions represented in TAM and the various individual differences, situational constraints, and managerially controllable interventions impinging on behavior" (Davis et al., p.988). Full mediation by beliefs and attitudes implies that all external variables would not exhibit any direct influence on usage intentions or usage behavior. Rather, such effects would only be exhibited indirectly through their relationship with beliefs. Thus, this study theorizes that individual differences exhibit an influence on the intention to use student performance information via their effects on an individual's beliefs about the new student performance information. This study will examine similar prior experiences and knowledge (McGeoch & Irion, 1952; Gick & Holyoak, 1987), the length of tenure in the workplace (Harrison & Rainer, 1992; Nickel & Pinto, 1986), and participation in training (Igbaria, Gamers, & Davis, 1995).

Perceived Information Quality

The perception of usefulness and easiness of a technology can be influenced by the quality of technology in itself as well as the user's individual differences. Empirically, the relationship between perceived output quality and perceived usefulness has been shown before (Davis et al., 1992). In the context of image theory (Beach and Mitchell, 1996, 1998), judgments of job relevance are

more apt to take the form of compatibility test whereby systems that are judged not be job-relevant are eliminated from one's choice set for further consideration. Judgments of output quality, in contrast, are less likely to be used for excluding options from consideration. Instead, they are more apt to take the form of a profitability test in which, given a choice set containing multiple relevant systems, one would be inclined to choose a system that delivers the highest output quality. An extended TAM (Venkatesh & Davis, 2000) posits that the perception of output quality is expect to explain significant unique variance in perceive usefulness. This study regards complexity and accessibility as the quality of information. The complexity of the testing information is defined as the level of sophistication of the information system in terms of its number of variables provided. Accessibility is defined as the level of acquisition of information by teachers to use the student performance data to set goals, measure progress, and discuss strategies.

CHAPTER THREE

METHODS

INTRODUCTION

The underlying purpose of this study is to explore those factors that affect teachers' intention to use of student performance information. The intention to use is measured though four variables: perceived usefulness, perceived ease of use, attitude, and external pressure (Davis, 1989). Perceived usefulness measures the extent to which a teacher believes that using student performance information will enhance his/her teaching performance, while perceived ease of use measures the degree to which the teacher believes that using student performance information will not require undue effort. Attitudes reflect an affective response to the behavior in question (i.e., a generalized liking or disliking for the behavior). External pressure is a construct measured by merging three components: pressure from peer teachers, campus administrators, and the central office. Chapter 3 describes the research questions, design, subjects, instrumentation, procedures, variables, and data analysis used in this research study.

RESEARCH QUESTIONS

The following questions guided this research:

1. What types of student performance information are available and accessible to teachers?
2. How do teachers use the testing information?
3. What factors influence teachers' use of student performance information?
4. Is the Technology Acceptance Model a useful conceptual tool for understanding the relationships among variables associated with teachers' intention to use student performance information?

STUDY DESIGN

The research questions lend themselves to an exploratory case study methodology (Merriam, 1998; Stake, 1995; Yin, 1994), nesting descriptive and inferential quantitative methods. This case study employed both survey research and individual focused interview methods. In other words, quantitative and qualitative methods were combined. The strengths of quantitative methods are that they produce factual, reliable outcome data that are usually generalizable to some larger population. The strengths of qualitative methods are that they generate rich, detailed, valid process data that usually leave the study participants' perspective intact (Denzin & Lincoln, 2000; Gall, Borg, & Gall, 1996; Patton, 1990).

With regard to incorporating quantitative and qualitative methods, though I recognize the argument that the purposes of the two methods are so different that

any reconciliation between them is bound to destroy the epistemological foundations of each (Rosenberg, 1988) or the paradigm of each, this study leans towards the tone that "… each method is based on different yet complimentary assumptions and each method has certain strengths that can be used to compensate for the limitation of the other" (Steckler, 1989, p. 115).

There are four ways that qualitative and quantitative methods might be incorporated; (1) qualitative methods are used initially to help develop quantitative measures; (2) qualitative methods are used to help explain quantitative findings; (3) quantitative methods are used to embellish a primarily qualitative study; and (4) qualitative and quantitative methods are used equally and parallel (Steckler, McLeroy, Goodman, Bird, & McCormic, 1992).

With emphasis on the second strategy, this study was predominantly quantitative. It examined the relationship among perception of usefulness, ease of use, attitudes, intention to use of student performance information, and external pressure to implement aligning tests, curriculum, and instruction by administering closed-ended questionnaire items and analyzing the responses (Research Question #4). The results of individual focused interviews and open-ended questionnaire items were used to help interpret and explain the quantitative findings (Research Questions #1, #2, & #3). One teacher from a selected school and a staff member from the central office were interviewed to learn about their strategies for

applying student performance information to instruction and/or campus improvement planning. Interviews probed for types of assessments, school culture regarding assessment, utilities, and professional development by adapting semi-structured interview protocols (Cromey & Hanson, 2000). In addition, written documentation was collected from the websites of the school district and the campuses that administered the survey questionnaires.

SITE SECTION AND SUBJECTS

Site Selection

This section describes the process of selecting a case in a qualitative way, and the process of selecting a sample from the population in a quantitative way. "Cases in qualitative research are selected by a *purposeful sampling* process (Gall, et al., 1996, p. 552)." Patton (1990) defines purposeful sampling as "the practice of selecting cases that are likely to be information-rich with respect to the purposes of a qualitative study (p. 169)." In this study, the general focus is the utilization of student performance information under the implementation of standards-based school reform at the school district level.

The school district was selected for the following reasons: (1) the long-term relationship with the Educational Productivity Council; and (2) their

demonstration of a typical response to the intensification of standards-based school reform in Texas.

The Educational Productivity Council, housed at the University of Texas-Austin, formed in 1980 with a mission to provide educational research designed to foster high performance standards for all students and accountability for responsible professionals in the K-16 system. In the 1980s, the EPC conducted productivity analyses to help member school districts target areas of resource inefficiency linked to underperformance. In the early 1990s, with the adoption of the Texas Assessment of Academic Skills (TAAS) test, the standardized exam used to assess Texas student learning, the EPC took on a new role of providing data support for districts and campuses. During this period, the EPC worked closely with member districts to transform testing data into informative district-, campus-, classroom-, and student-level analyses to help educators tailor instruction to student needs (Educational Productivity Council, 2002). The school district chosen in this study as a member district of EPC has been provided with the multi-level student performance analyses as well as training service for using them.

In the 2002-03 school year, the Texas Assessment of Knowledge and Skills (TAKS) will become the new statewide assessment program in 2003 for grades 3-11. This test is expected to be more rigorous than the Texas Assessment

of Academic Skills (TAAS) because it will be based on the state-mandated curriculum (the Texas Essential Knowledge and Skills - TEKS). The exit level test will be designed for 11th grade and will assess English language arts, mathematics, science, and social studies. The test will require knowledge of Algebra I, Geometry, Biology, integrated Chemistry and Physics, English III, and early American and United States History, World History and World Geography. The class of 2005 will be the first class required to pass the new assessment in order to graduate. The first year of implementation of the Student Success Initiative will be 2003. Students in grade 3 will be required to pass the state assessment in reading in order to be promoted to the next grade level without consultation of the grade placement committee. Beginning in 2005, students in grade 5 will be required to pass the state assessments in reading and mathematics in order to be promoted. Students in grade 8 will have to pass the state assessments in reading and math to be promoted in 2008 (TAAS Manager, n.d.; Texas Education Agency, n.d.;) .

The transition from TAAS to TAKS led the school district chosen in this study to develop its own database tools to effectively deal with student performance data, and to use the databases on each campus. In this transition period, it is becoming increasingly important for the central office of the school district to investigate how the use of benchmark tests, locally developed

assessments designed to assess TEKS mastery, can help teachers to better prepare instruction aligned with state-mandated curriculum and to participate in the campus improvement planning. These needs of the school district were consistent with the nature of this investigation as served as a primary reason the district was selected for this study.

The AA Independent School District (pseudonym) chosen for the study is a large suburban school district enrolling more than 33,000 students in central Texas. During the past five years, the number of students has increased by 2%. The district has a high percentage of white students and is predominately middle to upper-middle class socioeconomic status. More than 80% of the district's graduating seniors take the SAT and ACT college entrance exams, scoring well above state and national averages.

Subjects

Researchers use survey research to collect information on variables of interest by administering questionnaires to a sample from the population. The population to be studied in this investigation includes all the teachers of the district. At the time of the study, the school district employed approximately 2,300 teachers. Initially, 16 schools (9 elementary, 6 middle, and 1 high schools) were selected from 42 campuses in the school district for the study based on

nominations made by AAISD staff. Researcher provided two criteria for selection in order to obtain the representativeness from the population. The criteria were level of campuses and vertical team. The researcher considered the sample should be proportionably collected from elementary, middle, and high schools. The school district has four vertical teams. Each vertical team is composed of one high school and a couple of middle schools and elementary schools. The sample campuses were selected from each of the four vertical teams, which might be considered as "area probability sampling" (Fowler, 1993, p.20). The central office administrator also recommended putting into account the accountability rating of each campus. It would be a potential factor to impact teachers' perception about the benchmark tests and reports.

Visits were made to each of the selected campuses from March 2003 to September 2003. Campus visits were conducted at times scheduled and approved by the principal. Principals distributed the survey questionnaires to all faculty who potentially use AAISD Benchmark Assessment Reports. All participants were informed of the confidential nature of survey responses. They were also informed that their decision to participate in this study was voluntary and that they could decide not to participate at any time without any negative repercussions. Teachers were given about one week to complete a questionnaire and return it to the office. The researcher visited at least twice to encourage a high response rate.

Follow-up procedures (Dillman, 2000) for getting more responses were conducted in a limited way. After the first visit to collect questionnaires, an email note urging teachers to respond was sent to every teacher via the principal with an attachment of the questionnaire. Second visits were made one week later.

Of 680 surveys distributed, 223 teachers (32.8%) returned the questionnaire. In order to enhance the validity, however, the researcher decided to exclude 17 questionnaires which had over 20% missing data on the questionnaire. The final usable response rate was 30.3% (n=206).

As Cohen and Cohen (1983) suggested that 5% or even 10% missing data on a particular variable is not large, the missing data on the final usable responses (206) were replaced with an estimated score, the overall sample average. This method has the advantage of simplicity, but it is not sensitive to subjects' patterns of scores on other variables (Kline, 1998).

With 206 responses, the sample size is close to some recommended levels of 100 (proposed by Bollen, 1989), 150 (proposed by Boomsma, 1982), and over 200 (recommended by Anderson and Gerbing, 1982). But the sample size is lower than the recommended minimum level proposed by Hair (1992) and Klein (1998). Hair recommended, for multivariate analysis, that the sample size should be at least 5 times the number of parameters in the model. Since the hypothesized model included 57 parameters, the minimum response necessary would be (57*5)

or 285. The final usable responses (n=206) in this study is lower than this recommendation. The implication of not meeting the minimum sample size suggestions could be related to the instability of the statistical results. According to Kline (1998), "there is more sampling errors with small samples than with large samples. Thus, statistical estimates calculated within small samples tend to be less stable than within larger samples" (p. 43).

For confirming the results from the survey questionnaire, one central office administrator and one teacher from a middle school where the researcher conducted the survey were selected to be interviewed. The administrator from the central office served as director of accountability and audit department. He was actively engaged in developing the Benchmark Assessment System in the school district. The interview with him provided background information that contributed to the context of the study, including the purpose of the benchmark assessment as well as a detailed description of what the teachers are receiving in the Benchmark Assessment Reports. He also contributed to inform the content of the questionnaire by validating many of the concepts identified in the literature. The teacher who participated in the other interview has taught mathematics in 7^{th} grade for five years. She started teaching in the AAISD. The researcher selected her after meeting in a graduate class during summer of 2003 and for the reason

that she worked for the campus. She informed the researcher that she did not participate in the survey.

RIGHTS OF HUMAN SUBJECTS

The school district granted permission to conduct the study. A letter granting permission is found in Appendix A. Participants were asked to cooperate via Invitations to Participate in Survey and Consent Form to Participate Interview found in Appendix B and C. These letters described the study, the importance of each individual's participation, my role, followed by the request for participation and the guarantee of confidentiality.

OPERATIONALIZATION OF VARIABLES IN QUESTIONNAIRE

Individual Differences

The term individual differences can be interpreted most generally to connote dissimilarities among people including differences in perceptions and behaviors, traits and personality characteristics, and variables that connote differences attributable to circumstances such as education and experience. This study examined similar prior experiences and knowledge (McGeoch & Irion, 1952; Gick & Holyoak, 1987), the length of tenure in the workplace (Harrison & Rainer, 1992; Nickel & Pinto, 1986), and participation in training (Igbaria,

Gamers, & Davis, 1995). With regard to participation in training, this study asked about the degree of 'usefulness' of benchmark workbook (user guide) as a proxy of training variable because the school district did not provide any training session for using benchmark assessment reports to teachers.

Perceived Information Quality

An extended TAM (Venkatesh & Davis, 2000) posits that the perception of output quality is expected to explain significant unique variance in perceived usefulness. This study regarded complexity and accessibility as the quality of information. The complexity of the testing information was defined as the level of sophistication of the information system in terms of its number of variables provided. Accessibility was defined as the level of acquisition of information by teachers to use the student performance data to set goals, measure progress, and discuss strategies. The measurement of perceived information quality was adapted from Davis et al. (1992) and Venkatesh and Davis (2000). Respondents scored all items below on a 7-point Likert scale, with *Strongly Disagree* and *Strongly Agree* as the two endpoints and intervening points were labeled (see Appendix D: Questionnaire).

Items:

1. The Benchmark Assessment Reports are easily accessible.

2. The Benchmark Assessment Reports contain an appropriate amount of information.

3. The Benchmark Assessment Reports are of high quality.

Perceived Usefulness (PU)

Davis et al. (1989, p.985) define perceived usefulness as "the prospective user's subjective probability that using a specific application system will increase his or her job performance within an organizational context." In this definition, PU was linked to whether teachers will ultimately gain rewards from their use of a report including student performance information and ultimately the attitudes they have toward using the report.

PU variables were operationalized according to the recommendations made by Davis (1993) and Ajzen and Fishbein (1980). Respondents scored all items below on a 7-point Likert scale, with *Strongly Disagree* and *Strongly Agree* as the two endpoints. Davis (1993) reported the use of 10-item scales for the measurement of PU. In this study, five items were excluded from the set for the measurement of PU based on two criteria: either they were not relevant to this study through discussing among dissertation committee members and the central office personnel in the school district, or they were very similar to another item that had been included.

Items:

1. The Benchmark Assessment Reports help me to accomplish student analysis

2. The Benchmark Assessment Reports help to improve the quality of my work.

3. The Benchmark Assessment Reports improve my instruction of students.

4. The Benchmark Assessment Reports make it easy for me to make instructional decisions.

5. The Benchmark Assessment Reports are useful.

Perceived Ease of Use (PEU)

Davis et al. (1989, p.985) define perceived ease of use as "the degree to which the prospective users expect the target system to be free of effort." This concept checked whether it is easy for teachers to learn to interpret the student performance information, and to interact with other faculty members in a clear and understandable way. Davis (1993) reported the use of 10-item scales for the measurement of PEU. Agarwal and Prasad (1999) selected 4-item scales after verifying construct validity. In this study, three items were selected from the set for the measurement of PEU. Respondents scored all items below on a 7-point Likert scale, with *Strongly Disagree* and *Strongly Agree* as the two endpoints.

Items:

1. The Benchmark Assessment Reports are easy to use.

2. * The Benchmark Assessment Reports are difficult to use.

3. The Benchmark Assessment Reports are understandable.

* Reverse scaled item

Attitude

Attitude was measured using a four-item scale constructed according to the guidelines provided by Ajzen and Fishbein (1980). Respondents scored all items below on a 7-point Likert scale, with *Strongly Disagree* and *Strongly Agree* as the two endpoints.

Items:

1. The Benchmark Assessment Reports are fun to use.

2. *I dislike using the Benchmark Assessment Reports.

3. I like using the Benchmark Assessment Reports

4. The Benchmark Assessment Reports provide an attractive working environment.

* Reverse scaled item

Intentions to Use

Intention to use is the willingness of teachers to use the student performance information during work activities. The intention to use construct was measured using 4 item questions (Davis et al., 1989; Morse, 1999). Respondents scored all items below on a 7-point Likert scale, with *Strongly Disagree* and *Strongly Agree* as the two endpoints.

Items:

1. I intend to continue using the Benchmark Assessment Reports.

2. I intend to increase my use of the Benchmark Assessment Reports in the future.

3. I intend to use the information from the Benchmark Assessment Reports for problem solving and decision making

4. I intend to share data from the Benchmark Assessment Reports with other teachers.

External Pressure

This study reflects the impacts of three interrelated social forces, peer teachers, campus administrators, and the central office staff, impinging on an individual facing the opportunity to adopt or reject a new system. This construct is similar to subjective norm, voluntariness, and other social influence factors.

Subjective norms are defined as a "person's perception that most people who are important to him think he should or should not perform the behavior in questions" (Fishbein & Ajzen, 1975, p.302). Some researchers (Agarwal & Prasad, 1997; Hartwick & Barki, 1994; Moor & Benbasat, 1991) found that subjective norms had a significant effect on intention in mandatory settings but not in voluntary settings. Based on these findings, external pressure was measured by four-item scale. Respondents scored all items below on a 7-point Likert scale, with *Strongly Disagree* and *Strongly Agree* as the two endpoints.

Items:

1. My fellow teachers believe that I should use the Benchmark Assessment Reports.

2. * My use of the Benchmark Assessment Reports is voluntary.

3. My campus administrators believe that I should use the Benchmark Assessment Reports.

4. Central office personnel believe that I should use the Benchmark Assessment Reports.

* Reverse scaled item

INSTRUMENTATION

Questionnaire

The survey instrument consisted of three parts. Part I contained questions on Perceived Information Quality (PIQ), Perceived Usefulness (PU), Perceived Ease of Use (PEU), Attitude(At), Intention to Use (IU), and External Pressure(EP). Questions used to measure both PU and PEU were derived from the Technology Acceptance Model (Davis et al., 1989; Agarwal & Prasad, 1999).

PU received reliability scores of .92 (Davis et al., 1989), .93 (Chau, 1996), .95(Agarwal & Prasad, 1999), and .96 (Szajna, 1996). PEU also received reliability score of .87 (Agarwal & Prasad, 1999). Attitude received reliability score of .83 (Agarwal & Prasad, 1999). IU received reliability score of .60 (Agarwal & Prasad, 1999), .80 (Chau, 1996) and .90 (Davis et al., 1989). The following table showed reliabilities for the latent factors used in this study.

Table 1. Summary Statistics and Cronbach Alpha Coefficients for Latent Factors

Latent Factors	Number of Items	Cronbach Alpah Coefficient
Information Quality	3	0.74
Usefulness	5	0.95
Ease of Use	3	0.77
Attitude	4	0.81
Intention to Use	4	0.84
External Pressure	4	0.51

Part II requested demographic data on the following variables; age, gender, course subject, and level of campus. Then, individual differences were asked on the followings: similar prior experience, years of teaching, and user-guide rating. The first two variables measured the amount of experience and time. User-guide rating was considered as a categorical variable. Similar prior experience is assessed using three items: level of familiarity with personal computers, productive software, and internal email system, Lotus Notes. Respondents scored each item on a 7-point scale, with Never Used One and Used One A Lot as the two endpoints, and Used One Occasionally as the midpoint.

Part III contained three open-ended questions designed to gather additional descriptive data. The initial set of questions contributed to exploring more factors that may influence the use of student performance information. The questions were, "List the *positive* factors to impact on your use of the Benchmark Assessment Reports" and "List the *negative* factors to impact on your use of the Benchmark Assessment Reports" Then, the last question asked about the possible use behavior relating to student performance information. The question was "How do you actually use the Benchmark Assessment Reports?"

The questionnaire was developed after a thorough review of previous research that investigated the theory and practices of Technology Acceptance Model. At first, Section I measuring core constructs of Technology Acceptance

Model was composed of 30 questions. Three questions were adopted from Venkatesh and Davis (2000). Twenty two questions were taken from Davis (1993) and Agarwal and Prasad (1999), who studied the relationship between individual differences and the acceptance of new information technology. In addition, five questions were developed by researcher based on some findings from Agarwal and Prasad (1997), Hartwick and Barki (1994), and Moor and Benbasat (1991) who found that subjective norms had a significant effect on intention in mandatory settings but not in voluntary settings.

Then, the questions were modified according to input from dissertation committee members, two central office staff members, and three teachers in the school district. Based on their expert reviews, 23 question items were selected in the final questionnaire.

Interview

One teacher from one of the campuses where the survey administered and one staff member from the central office were asked more about the strategies for applying student performance information to instruction and/or campus improvement planning; types of assessments, school culture regarding assessment, utilities, and professional development by using semi-structured interview protocols (Cromey & Hanson, 2000).

The teacher and central office staff member were asked focused questions about the types of assessment reports. The questions also probed the subjectivity, method, purposes, and frequency about using the assessment results. Then, questions about the priority of learning about the assessment data and using it in their school and/or district were asked. The questions probed any formal or informal mechanisms in place at their campus to support the use of student assessment data, including professional development opportunities to support their use of student assessment information. Finally, the interviews investigated the alignment among test results, curriculum, and instruction, and the contribution of student assessment information to campus improvement planning process.

The interview ranged from 30 minutes to 40 minutes. Each interview was audio taped and transcribed. These data provided me with information primarily directed toward answering research questions # 1, #2, & #3 (types of student performance information available and accessible to teachers, their actual usage of student performance information, and organizational factors to impact on the use of student performance information).

Written Documentation

The staff member from the central office was asked to provide examples of written documentation that would provide concrete evidence of their

implementation of the Benchmark Assessment Repots. This documentation included the memorandum from the central office to a local campus, examples of the Benchmark Assessment Reports, a user guide for the benchmark test, the district's testing calendar, and other testing information. These materials were analyzed to explore the type of student performance information.

ANALYTICAL PROCEDURES

I conducted interviews to collect additional data for Research Question #1, #2, and #3 (types of student performance information available and accessible to teachers, their actual usage of student performance information, and organizational factors to impact on their use of student performance information). The interviews served both confirmatory and exploratory purposes.

In addition, the text responses to open-ended questions about additional factors that impact the intention to use of student assessment information positively and negatively and actual usage of the information were captured for answering research question #2 and #3

I used Microsoft Word to transcribe the interviews and responses from open-ended questions in the survey questionnaires and saved each of those data as a text file. All files were imported into *QSR NUD*IST*, a computer program designed for qualitative analysis methods.

I followed the procedure outlined by Creswell (1998).

- Open Coding. The first phase of data analysis involved a search through the entire data set and assignment of codes.
- Category Formation. The hierarchical category definitions reflected the specific events that occurred and influences of these events. Each was summarized in the form of answers to each research question.
- Case Comparison. I compared the themes that emerged from each of the cases. This analysis looked for similarities and differences in the interview and responses from open-ended questions.

Testing procedures relating Research Question #4 (testing which factors have an influence on the intention to use of student performance information based on the theory of Technology Acceptance Model) for this study involved Structural Equation Model (SEM), a second-generation multivariate technique (Bollen, 1989) which has gained popularity in use among published Managing Information Science (MIS) studies (Chau, 1996). SEM is a comprehensive, flexible approach to modeling relations among variables. The primary aim of SEM is to model covariances, which entails proposing a set of relations (i.e., a model) and evaluating their consistency with the relations manifest in an observed covariance matrix (Bollen, 1989). The general structural equation model consists

of two complementary models: the measurement model, of which factor analysis is an example, and the structural model, which concerns relations among independent and dependent variables and of which general liner modeling is an example.

Each of these components was analyzed on the following two steps. First, the measurement model was tested for convergent and discriminant validity. The initial measurement model imposed a model where all factors were allowed to covary. If the initial measurement model does not fit satisfactorily, the measurement model must then be adjusted and fixed prior to examination of the structural model (Segars & Grover, 1993; Stapleton, 2001). At this step, confirmatory factor analysis (CFA) and reliability analysis via EQS (Bentler, 1995) were performed on the 23 manifest variables representing the 6 latent constructs.

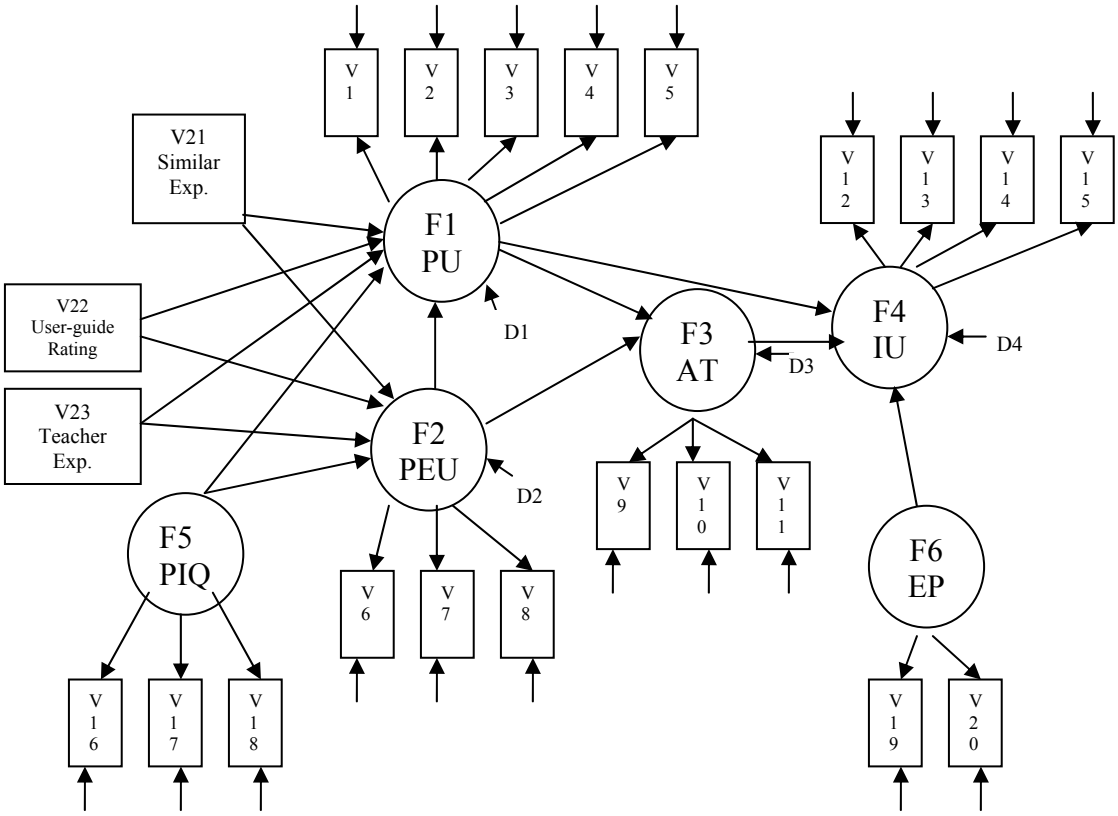


Figure 4: Measurement and Structural Model

F1 – F6: Factor 1 – Factor 6	PU: Perceived usefulness
PEU: Perceived ease of use	AT: Attitude
IU: Intention to use	PIQ: Perceived information quality
EP: External Pressure	V1-V23: Manifest variables 1-23
D1-D4: Error of dependent factors	

Second, the initial structural model imposed on the latent factors defined in the final measurement model was then tested for explanatory power and goodness of fit. The model was further tested for goodness of fit and modified again so that the "best fitting" model could be used for the final tests. Using EQS

(Bentler, 1995), a structural equation model was developed and testing procedures were employed using Maximum Likelihood estimation procedures.

CHAPTER FOUR

RESULTS

A purpose of this study was to examine (1) what types of student performance information are available to teachers; (2) how they use the information; and (3) what factors impact positively and/or negatively on teachers' use of student performance information. This study also examined (4) the relationship among perception of usefulness, ease of use, attitude, intention to use student performance information, and external pressure to use the information.

In the following sections, I presented a preliminary analysis based on descriptive statistics to provide 206 teachers' characteristics in the survey administration. Then, I provided four major sections which pertain to the analysis of research questions #1 through #4 respectively.

Using the responses from three open-ended questions in the survey questionnaire and two interviews, the analyses and findings on research question #1 to #3 were reported first. The teacher and central office administrator completed a consent form to participate (Appendix C) and participated in a one-to-one interview lasting approximately half an hour. Each interview was audio taped and transcribed. I reviewed, coded, and analyzed the transcripts.

Then, the results of the statistical analyses were used to address the research question #4 and the statistical findings for each hypothesis.

DESCRIPTION OF DEMOGRAPHIC DATA

Preliminary analysis included the descriptive statistics to provide teachers' characteristics, such as age, gender, course subject, and grade level. Then, individual differences were asked on the following: year of teaching, the degree of usefulness of the Benchmark Workbook as a user's guide, and prior similar experiences of the local school district email system, productive software and personal computers in general.

The descriptive statistics are shown in Table 2. There were 206 teachers from an suburban school district in Texas (89 in nine elementary schools, 94 in six middle schools, and 23 in a high school) participating in the study. 84.5% of the participants were female and 12.1% were male with 7 respondents who did not specify gender. The gender distribution of the sample was similar to that of the school district. Of the total 2,319 teachers in the school district, 1,912 (82.4%) were female; and 407 (17.6%) were male.

Teachers in the sample had spent an average of 12.5 years in education and 7.4 years in the school district, whereas teachers in the population have taught an average of 10.8 years in education and 5.8 years in the school district. This

study might have more responses from the teachers who have more teaching experience both in education and in the school district.

A 32% response rate was not too strong. The most severe problem this study might have to face was non-response bias that the sample generated might not be representative of the population. The characteristics of those who responded to the survey questionnaire could be different from the characteristics of those who do not respond. That is why the comparison of gender and teaching experiences between the sample and the population was made in this section. Tests for non-response bias were conducted comparing the distribution of gender and teaching experience in education, and the means of the teaching experience in the school district. The test revealed no significant difference in gender between the sample and the population (X^2 = 3.51, df=1, p=.061). However, the tests revealed significant difference in teaching experience in education (X^2 = 15.96, df=3, p=.001), and in teaching experience in the school district (t=3.64, p<.001). Sample respondents were more experience in teaching than the overall population.

A mean value for three items, capturing a respondent's prior similar experiences with the district email system, software, and personal computer in general was used in the analysis. The data show that teachers in this sample had a high use of the computer facilities relating to the student performance

information. With a regard to using a user's guide for the benchmark assessment report, only 12.1% of respondents indicate that the guide is useful. Most teachers (83.5%) report that they have never seen the guide.

Table 2. Presentation of Demographic Data (206 Teachers)

Gender	
Female	174 (84.5%)
Male	25 (12.1%)
No Response	7 (3.4 %)
School	
Elementary	89 (from 9 campuses)
Middle	94 (from 6 campuses)
High	23 (from 1 campus)
Age	
21 – 30	44 (21.4%)
31 – 40	56 (27.2%)
41 – 50	51 (24.8%)
51 or more	43 (20.9%)
No Response	12 (5.8%)
Average Teaching Experiences	12.5 years (SD=6.63)
Average Teaching Experiences in AAISD	7.4 years (SD=8.45)
Average of Prior, Similar Experiences	6.76 (SD=.65)
(Lotus Notes, Productive Software, & Computer)	
User Guide	
Useful	25 (12.1%)
Not Useful	9 (4.4%)
Never Seen it	172 (83.5%)

RESEARCH QUESTION #1

What type of student performance information is available and accessible to teachers?

The findings presented in this section are the result of teacher and central office administrator interviews and document review data obtained from AAISD (pseudonym) and the websites of sixteen schools. The results are presented according to three different levels of student performance tests-state, district, campus and classroom.

State Level of Tests

Texas Assessment of Knowledge and Skills (TAKS): As mandated by the Texas Legislature in 1999, the Texas Assessment of Knowledge and Skills (TAKS) replaces the Texas Assessment of Academic Skills (TAAS) and was administered for the first time in the 2002-03 school year. The TAKS measures the statewide curriculum in reading at Grades 3 – 9; in writing at Grades 4 and 7; in English Language Arts at Grades 10 and 11; in mathematics at Grades 3 – 11; in science at Grades 5, 10, and 11; and social studies at Grades 8, 10, and 11. The Spanish TAKS is administered at Grades 3 through 6. Satisfactory performance on the TAKS assessment in English, mathematics, science, and social studies at Grade 11 is required to receive a high school diploma.

Reading Proficiency Tests in English (RPTE): The RPTE was field tested in the spring of 1999 and implemented in the spring of 2000. These tests are designed to measure annual growth in the English reading proficiency of second language learners, and are used along with English and Spanish TAKS to provide a comprehensive assessment system for limited English proficient (LEP) students. LEP students in Grades 3 – 12 are required to take the RPTE until they achieve a rating of "advanced".

State Developed Alternative Assessment (SDAA): SDAA assesses special education students in Grades 3 – 8 who are receiving instruction in the Texas Essential Knowledge and Skills (TEKS) but for whom TAKS is an inappropriate measure of their academic progress. This test assesses the areas of reading, writing, and mathematics. Students will be assessed at their appropriate instructional levels, as determined by their admission, review, and dismissal (ARD) committees, rather than at their assigned grade level. The SDAA is administered on the same schedule as TAKS and is designed to measure annual growth based on appropriate expectations for each student as decided by the student's ARD committee.

District Level of Tests

Benchmark Assessment: To measure student progress throughout the year and predict their TAKS readiness, AAISD implemented a series of "benchmark assessments" for grades K-12 in reading, writing, mathematics, science, and social studies. These Benchmark exams are scheduled to be administered two or three times in each grade and each subject in a year.

Aside from the locally developed benchmark assessment, AAISD administers other tests, such as the OLSAT, DAT, and Gates-MacGinitie. The **OLSAT** is a test of aptitude of school abilities index. It is administered in grades 2, 4, 6, and 8. The **DAT** test is administered at the 8th grade level. It is a career interest survey test and it is not related to the benchmark assessment. **Gates-MacGinitie** is administered in grades 6, 7, and 8. It is a reading test for the approximate reading level of school children. It is used for placement of students in special classes at the high school level and at the middle school level, such as Pre-AP class and Honors class.

In order to ensure that students in the district were progressing sufficiently to pass the Grade 3 reading test in 2003, AAISD developed a Kindergarten Reading Assessment (KRDA) and District Reading Assessment (DRA). These assessments were first used with Kindergarten students in 1999, and successive grades were added each year. These assessments determine which students are

reading above, on, or below grade level, and provide data to identify children that are struggling with reading so they may receive intense, accelerated instruction.

Campus and Classroom Level of Tests

At the local campus and classroom levels, there were many tests based on the diversity of student groups, program implementation, and subjects. One staff member in the central office remarked,

> There are other tests that are administered at local campuses. Some may be for special populations. We do have a Metropolitan Achievement Test that is administered to a special program of children who speak other languages. We are required by the state to administer those tests to children LEP. And there may be some special ones for special education better specific to child with certain handicaps.

The current assessment practices at the classroom level were reportedly based on individual teacher interpretation of the state and district standards and curriculum (TEKS in state and scope and sequence in AAISD) and a variety of data collection tools. The teacher reported,

> In my classroom, I used data from teacher made tests that I have made for the end of units and any quizzes that I have made. And I also look at the homework grades to assess how students are learning if they are progressing with the curriculum. If they are not, then what intervention needs to be made. As a campus, I would say that every teacher looks at the same data. We all look at our tests we have made to give to the kids, and look at quiz grades. We also of course look at TAKS data from last year from the spring administration. We do look at benchmark data. We have not had the first benchmark yet from math release. But once that happens

> we will be looking at that data to see what we can do. … I would say 12-15 teachers made tests for 7th grade math during one school-year.

In summary, AAISD with well-developed assessment systems (Benchmark Assessment, KRDA, DRA, OLSAT, DAT, and Gates-MacGinitie) responded in an organized way to rising standards by improving alignment within their local curriculum (AAISD Scope and Sequence) while ensuring that local standards were consistent with those of the state curriculum (TEKS) and state mandated tests (TAKS, RPTE, and SDAA). Schools and classrooms also were engaged in ongoing initiatives to improve student achievement under their own circumstances.

Benchmark Assessment Reports

This section provided a brief explanation about what the teacher were receiving in the Benchmark Assessment Reports in order to fully comprehend the findings on Research Questions #2, #3, and #4.

The detailed analysis was an item analysis regarding how students performed on the test showing the number and choices of response given to each question. Individual teachers who have an access to manipulate the data should be able to pinpoint for every student's individual responses. The reports included each individual teacher's reports, their classroom reports, and their campus

reports. The reports gave a variety of demographic information about students' ethnicity, economic disadvantage, special education, gender, etc.

The campus reports analyzed the results based upon the standards that the state has for passing. The school district also has a local standard for passing and another higher standard, which is called 'proficient level'. For example, the state may have a 60% passing rate on a test item. The school district may set the local standard for passing at 70%, and the proficient level at 85% of items correct. The following reports consisted of the Benchmark Assessment Reports for 6th, 7th, and 8th Grade Writing for a middle school (see Appendix H).

- Campus Summary Reports
    - Average number of percent of total items correct
    - Average percent mastery by objective
    - Percent Passing at standard level (70% or higher) and proficient level (85% or higher)
- Campus Item Analysis
    - Percent of responses per test item
    - Matching of test items to TEKS and TAKS objectives
- Teacher Item Analysis
    - Percent of responses per test item
    - Matching of test items to TEKS and TAKS objectives
- Teacher Reports
    - Number and percent of total items correct by each individual student (sorted by teacher and course code)

    - Mastery and number of items correct by objective by student (sorted by teacher and course code)
- Flexible Campus Level Generated Reports
    - Excel spreadsheet sent with report will allow for further analysis of campus data via sorting by ethnicity, special education, economically disadvantaged, percent correct and mastery levels.

RESEARCH QUESTION #2

How do teachers use the testing information?

The findings presented in this section are the result of interviews and the responses from one of the open-ended questions in the survey questionnaires. The survey collected qualitative data through three open-ended questions in hopes of gathering rich descriptive information either to confirm responses to close-ended questions or to explore themes not easily formatted into close-ended questions.

Below are descriptions of the response characteristics from qualitative data collected for the last questions.

Table 3. Teachers' Usage of AAISD Benchmark Assessment Reports

Factor	Number of responses
Adjustment of instruction	31
Tutoring on weak items	16
Don't use/Minimum use	10
Meeting for planning	5
Preparation for TAKS	5
Comparison among class	5
Others	25

Ninety-seven teachers provided narrative responses. The objective for this question was to gather more descriptive data concerning the diversity of teachers' actual use of student performance information. Because the survey questionnaire asked about teachers' intention to use student performance information, this open-ended question about teachers' actual use of the information would provide sufficient examples in teachers' real lives.

In the largest category, 31 teachers submitted responses suggesting that they used the benchmark assessment information to check the efficacy of local curriculum and instructional practices. Changes to instructional practices occurred at school and classroom levels and at the individual student level. Teachers provided a flexible approach to their teaching strategies based on students needs. Examples of these responses are:

- I look at the percent that my students got correct of the TEKS I have actually taught. I want to make up my own Benchmark test. Then I can actually use the results as a valid assessment for my students. This was a bank of questions for each TEKS and TAKS objectives and I will make the tests to reflect what I have actually taught.

- I record the question number missed on the Benchmark and the objective it tested. Then, I can determine if it is a problem with my teaching, something that hasn't been taught, or if it is an individual student that needs reteaching.

- I check to see which objectives we scored poorly on and from there I've tried to devise a plan of how I can improve my instruction based on the low scoring objectives.

- Gives ideas on instructional differentiation with certain students.

- The reading benchmark will help me pick the lowest areas to work on, and the writing helped me to see that the students are not making much process writing in 5th grade. A majority of my kids were not familiar with 6th traits.

When the teacher who participated in the interview was asked to describe how she actually uses the benchmark assessment report, she stated that she would reteach based on item analysis of the benchmark test, examining response patterns on specific test items.

> But we look at by objective how each of our kids did, and then we go to do item analysis. … If only 55% of my kids knew C was the correct answer, then what I need to go back in and reteach that. And we also then go to each individual student. We can look at each individual student on how they did on each objective and on each question. … It's real easy to look at overall this is how my kids did on objective 1, 2, 3, 4, and 5. It's

> much easier to do that, then just say this is how this kid did on this objective, this kid, … we have got 120 kids. So it is a lot easier to look at comprehensive is group then each individual student.

Sixteen teachers reported that they used the test results for identifying students' weakness and providing tutoring. Ten teachers also responded in a similar way; they focused on assessing TEKS and/or objective mastery. For example:

- We drilled the weak item analysis areas in warm ups. The writing scores were bad-We did an intense unit on how to respond to these better.

- Students who failed benchmarks were required to attend tutoring with their teacher and focus on objectives/questions missed. This helped to recognize who needed what help.

- We will be forming tutoring groups and comparing which teaching delivery methods are producing results.

- To see if TEKS is understood by students.

- Note student's mastery by objective.

- Mainly to determine need to re-teach or need to re-evaluate how a TEKS is addressed in class.

The interviewee also described tutoring as a strategy to cover students' weakness in their performance.

> Once a week, we are actually pulling students in from one of the elective classes, like from PE or from Band or from Art or dance. One of those elective classes, just once a week we are pulling him in from elective classes, and working with him in class period. So I think that is really helpful.

Ten teachers reported they didn't use the benchmark assessment reports, or used them at the minimum level as directed by the school district. The dominant sub-category for these responses was the dislike of frequent tests. In the next sub category teachers commented about no need for using the reports because they already knew the results. They stated:

- They don't give me anymore information than I already get from my test.
- I file them. By the time I get them, I have already graded them and gone over them with students.
- I don't use the data. I know my students without the assessment. The data does not add to my knowledge.
- I do them because the district wants us to. I don't use them for evaluating students.
- I use them because we have to.

Several other categories, although they are limited in number, also provide various ways teachers' use of teachers' student performance information. Five teachers reported that they used the reports for planning in conference and staff

meetings. Another 5 teachers stated that the reports were used for preparation for the state mandated test, TAKS. Furthermore, responses showed that the reports were used for comparing classes (5 responses), examining test quality and relationship among variables (2 responses), special education students (3 responses), and other. For examples:

- Discuss results with team to prepare students better.
- I use it as benchmark of TAKS for 5th graders entering 6th grade science.
- Also look at low scores and focus on students who need extra guidance to do well on TAKS.
- See how my students are doing compared to rest of students at this school and in the district.
- I read them to see how my students did, particularly my learning disabled or special education students.
- I use it to find out the correlation between the success of lower socioeconomic students vs. higher economic.
- I do them and look over the results to see if the rest of my assessments (that are much more useful to my students) seem to say the same thing.

The interviewed teacher also reported that she has made positive methodological changes in response to comparing with other teachers' performance.

> When we get the report back, we look at first of all overall how our students did, and then basically we got to a breakdown if there are four seventh grade math teachers. We've got a breakdown and it shows on one line here is how teacher A's kids did, here is how teacher B, teacher C, teacher D, and then what able look at and say, OK your kids are whole a lot better in object 1 than mine. How did you teach what we were doing different than me? And a lot happens, because we plan together. But obviously the way you are delivering instruction is little bit different.

In summary, teachers used the benchmark assessment information (1) to check the efficacy of local curriculum and instructional practices; (2) to assess state curriculum standards and/or objective mastery, and (3) to prepare state mandated tests. However, the study also found that some teachers did not use the benchmark assessment reports, or used them at the minimum level as directed by the school district.

RESEARCH QUESTION #3

What factors influence teachers' use of student performance information?

Table 4. Positive factors to impact on your use of AAISD Benchmark Assessment Reports.

Factor	Number of responses
Identification of student needs	23
Alignment of curriculum and tests	14
Preparation for instruction	12
Information access,	7
Information quality	5
Others	31

For this question, 92 teachers provided narrative responses. After coding and transforming the responses, the results were categorized into the following themes: identification of student needs on student performance, alignment of curriculum and the test, preparation for instruction, information access, information quality, and others.

Twenty-three teachers responded that efforts to identify student needs positively impacted their use of benchmark assessment reports. For example:

- We can more accurately pinpoint the kid's strengths and weaknesses.
- The data is individualized for each student.
- The writing benchmark used at the beginning of the year helped me to get an understanding where my students in their writing skills.

Fourteen teachers responded that alignment of curriculum to the test provided them with more intention to use the student performance information. For example:

- They could be useful if they match the TAKS.
- They keep us (teachers) striving for TEKS mastery.
- It is good to promote a district expectation of following the TEKS.

Responses concerning preparation for instruction were submitted by 12 teachers. Preparation for instruction categories were created based on responses concerning belief on usefulness of data, alignment of instruction and test results, and willingness to re-teach. Participants reported:

- I found it helpful in preparing and planning focused instruction
- It helped me to know what should be stressed or retaught.
- I enjoy seeing what information my students retain. I also use the information to teach, or re-do lessons on objectives that students score low on.

Seven teachers stated that timely return of the reports would have helped them to use the information. These respondents stated:

- If provided feedback quickly, I could compare student achievements on the 2 benchmarks.
- In a perfect world they should give you information about your students. This year is not returned to teachers quickly enough to be useful.
- Get them back-quick turnaround.

Five teachers claimed that the stronger reliability and validity of the benchmark test would lead them to use the reports. They reported:

- It would be great if we feel the test itself was a real assessment of what students know, especially when comparing results for improvement.

- If the quality of the tests improves, obviously benchmarking can provide information to make curriculum decisions and interventions.

- It could be useful when mistakes are gone.

In summary, teachers reported identification of students' need, alignment curriculum and tests, preparation for instruction, information access, and information quality positively impacted their use of benchmark assessment information.

Table 5. Negative factors to impact on your use of AAISD Benchmark Assessment Reports.

Factor	Number of responses
Poor quality of test and reports	20
User unfriendly format	17
Low information access	16
Time consumption	16
External pressure	9
Misalignment with curriculum	9
Others	12

To this question 97 teachers provided narrative responses. After coding and transforming the responses, the results were categorized into the following themes: poor quality of benchmark test and its report, user unfriendly format, low

information access, time consumption, external pressure, misalignment with curriculum, and others.

In the largest category, 20 teachers responded that the poor quality of benchmark tests negatively impacted their use of the benchmark assessment reports. Validity and reliability of the test, relevance and accuracy of test, and amount of information were the three dominant sub-categories. For example:

- This data is not relevant to students at the level I teach.
- Many of the tests were invalid due to wrong answers; more than one correct answer; topics questioned were not on the TAKS.
- Poor test quality, inaccurate/incorrect data
- The reports give me the results of test that are not appropriate either because of level of difficulty, questions for a different grade level, or do not match the TEKS.
- The wording of questions and the answer choices for the reading benchmark were too elusive-our entire team together could not even figure out the correct answers. The math benchmark seemed like it was more focused on "tricking" the students rather than truly assessing.
- They are not always developmentally appropriate at assessing the TEKS. Also, some of the data has returned inaccurately.
- They were frequently inaccurate and the quality of the questions (particularly language arts) made it impossible to draw conclusions from results.

- Not enough questions over each TEKS to provide an accurate mastery/passing level. For example, there will be 2 questions over a TEKS, so if a student misses one question, they will score 50%.

- So much information is printed on the reports. It is overwhelming trying to find individual students' data

Seventeen teachers claimed that user unfriendly format negatively impacted their use of the reports. User unfriendly format was sub-categorized into not easy to use, hard to understand, and not individualized results. Participants stated:

- The benchmark results seem "user unfriendly"-very technical and hard to read.

- I find it confusing to determine specific areas of difficulty for students.

- Reports are not easy to read.

- It basically is hard to use on the computer. The school should print out copies and have one per team. I don't have time to fiddle with spread sheets.

Sixteen teachers submitted responses concerning information access. This concern mostly meant late returns of the test results. These respondents stated:

- Response time is so slow. I usually had grade and have kids graph results before they are sent to central office.

- We don't receive the results in a timely manner.

- Teaching time is used to administer so many tests. Results are not shared in a timely manner that could be used to direct our instruction-still waiting for January results (March).

- I usually get results after I've scored them myself.

- Too slow in delivery of data. Too long after the testing date. Difficult to access

Sixteen teachers stated that taking a lot of instruction time negatively impacted their using benchmark assessment results. They stated:

- The tests take a lot of instructional time.

- Takes away too many instructional days. We never should have used instructional time for all of us to administer. I lost instructional time in my class which actually is held accountable on the official TAKS test given in April.

- Teaching time is used to administer so many tests.

- I know we lose instruction time due to having to assess our students per grading period.

- It disrupts my lesson plans and ability to complete semester goals.

The interviewed teacher remarked that the lack of time to review the benchmark assessment reports negatively impacted on the teachers' intention to use student performance information.

> Time is the biggest obstacle, biggest hindrance, I would say. Because we are not given time during our work week per se, that says OK teachers we are giving you this time to look at your data. When we look at our data, it is coming out of our own time. Before we get there, before 7:45am before our clock time and then after our clock time at 3:45 pm.

Nine teachers stated the pressure and stress of high stakes testing negatively impacted their using students' performance information. For example:

- "High stakes" testing is not useful in any way
- Way too much pressure for students and teachers.
- Too many benchmarks cause students and teachers a lot of stress.
- We are asked to do them. No one in administration seems to care.
- District "mandates" about use.
- So much emphasis on them by district. DATA, DATA, DATA !!! Useful, but overstressed.

The staff member from the central office and the teacher who participated in the interview shared the amount of stress that 5th grade teachers are going to

have next year due to intensification of accountability. They also identified the extent to which teachers felt overwhelmed by a growing number of assessments might be dependent on the subject.

> In 2005, if they don't pass the test in 5th grade, they will be retained. And then 2008, if they don't pass in 8th grade, they would be retained. So, for example, my campus we are looking OK. It's 2003 now. It's almost 2004. 2008 is really not that far away. So we are really looking at four, little over four years, we are going to be dealing with this fact. We may have a lot of 8th graders retained if they don't pass the test. Yes, I would say there is definitely pressure in the middle school too.

> I think everybody feels the pressure differently. I think that the math teachers, the reading teachers, the social studies, science teachers probably feel the pressure a little bit more than maybe the PE teachers or some of the elective teachers. And I think they feel the pressure too. But I don't think it is the same. When your scores were posted in front of the whole entire school, and then shows you know the math scores, the reading scores, the social studies, and science scores. If you are not one of those teachers, you don't have a real connection to it. So, I think pressures are little bit heavier for those teachers.

Other teachers stated misalignment with curriculum, evaluation of teachers, and preparation just for state mandated test negatively impacted their using benchmark assessment results. They stated:

- Not aligned to what we have gone over. There is variation in scope and sequence and we have not always covered what was being tested.

- The test does not test what I have taught because I can't cover the amount of content in the prescribed time and I have to teach other things.

- Reports were used to compare schools and teachers.

- We were told that they are just for our information- Why do I have to send it in to the district-just for comparison of school/teachers?

In summary, teachers reported poor quality of test and reports, user unfriendly format, low information access, time consumption, external pressure, and misalignment with curriculum negatively impacted on their use of the benchmark assessment information.

RESEARCH QUESTION #4

Is the Technology Acceptance Model a useful conceptual tool for understanding the relationships among variables associated with teachers' intention to use student performance information?

To determine the effect of perceived usefulness, perceived ease of use, perceived information quality, external pressure, and the attitude on intention to use student performance information, the data were applied to a structural equation model and analyzed via the EQS program described in Bentler (1995). The model was run in two phases (1) the measurement model and (2) the structural model.

Using EQS, a structural equation model was developed and testing procedures were employed using Maximum Likelihood (ML) estimation procedures. ML estimation was the default method in the EQS. "Maximum

likelihood describes the statistical principle that underlies their derivation: if they (the estimates) are assumed to be population values, they are ones that maximize the likelihood (probability) that the data (the observed covariances) were drawn from this population (Kline, 1998, p. 125)." ML estimation assumes multivariate normality, which means: all the univariate distributions are normal; the joint distributions of any combination of the variable are also normal; and all bivariate scatterplots are liner and homoscedastic (Kline, 1998). This study ran EQS not with raw data but with covariance matrix, so the values of indexes of multivariate skew or kurtosis by Mardia (1970) could not be produced. Therefore, multivariate normality was checked through the inspection of univariate distributions. The data sets with absolute values of univariate skewness and kurtosis indexes seemed to be described as normal (see Appendix F).

Measurement Model

The first step in the test was to determine whether the manifest variables were loaded on latent factors appropriately on which the Technology Acceptance Model is based. This step was accomplished via confirmatory factor analysis (CFA) using the EQS program. Generally, CFA is viewed as the first step in developing a structural equation model, as it allows for hypothesized relationships

between manifest variables and latent factors to be tested statistically (Kline, 1998).

The CFA using the EQS program focused on 6 latent factors (perceived usefulness, perceived ease of use, attitude, intention of use, external pressure, and perceived information quality) and 23 manifest variables. The factor structure revealed high and consistent loadings on the six proposed factors. Table 6 shows statistics regarding initial item reliability and construct reliability.

Item reliability refers to the degree of variance explained by the construct rather than by error. This is typically measured by squared factor loadings which represent the item ability to capture variance within the construct. Items demonstrating high reliability typically record squared factor loadings more than .50 as recommend by Fornell and Larcker (1981).

Variance Extracted refers to the amount of variance in indicator variables for a factor captured by the measurement model versus the amount due to measurement error. Variance Extracted can be calculated by (summation of squared factor loadings)/ number of variables. It has been suggested that Variance Extracted should be greater than .50 to demonstrate significant variance captured by the measurement model (Fornell & Larker, 1981).

Construct reliability "assesses whether a measure relates to other observed variables in a way that is consistent with theoretically derived predictions (Bollen,

1989, p.188).” Traditionally, construct reliability can be calculated by the summation of squared factor loadings / [(summation of squared factor loadings) + (summation of error variance)]. A value of .70 or greater suggests evidence of strong construct reliability (Ruth, 2000). This study used a proposed improved index of construct reliability, Coefficient H, using squared factor loadings *l*;

$$H=1/(1+(1/[l^2/(1-l^2)+\ldots+l^2/(1-l^2)]))$$

Recommended minimum size of H is .70 to .80 (Hancock & Mueller, 2001). It has been argued that the traditional construct reliability is adversely affected if it has some indicators loading with opposite sign, if an additional (relatively poor) indicator detracts from the overall assessment of construct reliability, and if the assessment of construct reliability is less than that of its single best indicator. However, the coefficient H is not affected by loading sign, not detracted by additional indicators, and not smaller than the reliability of the best indicator.

Table 6. Initial Test of Reliabilities and Validities

Item	Label	Squared Factor Loadings	Variance Extracted	Construct Validity (H)	Disposition
	Perceived Usefulness		.782	.906	
PU1	Student Analysis	.698			Retain
PU2	Quality Work	.775			Retain
PU3	Instruction	.859			Retain
PU4	Decision	.808			Retain
PU5	Useful	.768			Retain
	Perceived Ease of Use		.554	.652	
PEU1	Ease	.760			Retain
PEU2	Difficult	.474			Retain
PEU3	Understand	.428			Retain
	Attitude		.532	.803	
AT1	Fun	.425			Retain
AT2	Dislike	.594			Retain
AT3	Like	.875			Retain
AT4	Attractive	**.232**			**Remove**
	Intention to Use		.585	.704	
IU1	Continue	.660			Retain
IU2	Increase	.674			Retain
IU3	Problem Solving	.609			Retain
IU4	Share Info	.396			Retain
	Information Quality		.534	.856	
IQ1	Accessible	.340			Retain
IQ2	Amount	.922			Retain
IQ3	Quality	.340			Retain
	External Pressure		**.381**	N/A	
EP1	From Teacher	**.168**			**Remove**
EP2	Voluntary	**.040**			**Remove**
EP3	From Administrator	1.000			Retain
EP4	From Central Office	.314			Retain

Three items that had squared multiple correlations with the latent factors of less than .30 and variance extracted and construct validity below recommended minimum level, .50 and .70 to .80 respectively, were dropped from the analysis in an attempt to explain more variance and aid in construct parsimony.

The tests show a considerably better measurement model with even more parsimony. The revised measurement model contains 20 manifest variables representing 6 latent factors. Squared factor loadings range from a low of 0.360 to a high of 0.829. Most loading values are above recommended minimums. Variance extracted ranges from 0.504 to 0.783 all exceeding the suggested minimum of .50.

Before running the CFA, the researcher set a single loading for each factor to equal one. The model converged in nine iterations, and results indicated an adequate fit of the measurement model. Although the overall X^2 statistic for model fit was significant ($X^2_{(df)} = 366.44_{(197)}$, $p<.001$), all other indicators of fit were within appropriate values. Because of X^2 is quite sensitive to the large sample size (Bentler, 1980; Bentler & Bonnet, 1980; Newcomb, 1990). Therefore, other indicators should be examined for the model fit (Bollen, 1989; Taylor & Todd, 1995). The Comparative Fit Index (CFI) was .940, where .950 is generally considered to be a reasonable cut-off for model fit. Another fit indicator that has gained acceptance for structural equation models, was at the acceptable limit (RMSEA<.10). Recently Hu and Bentler (1999) suggested two combinational rules: a cutoff value of .96 for CFI in combination with SRMR>.09 resulting in the least sum of Type I and Type II error rates; RMSEA>.06 and SRMR>.09. According to Hu and Bentler (1999), "when N<250, the recommended

combinational rules are more preferable because the rules tend to reject more simple and complex true-population models under the nonrobustness condition (p.28)." Based on this joint criteria, the final CFA met the cutoff of fit indices.

Additionally, the Lagrange Multiplier modification index, which approximates the amount by which the model's overall chi-square would decrease if a particular parameter is allowed to be loaded freely onto any factor, revealed that six additional covariance of residuals for (Quality Work – Useful), (Understand-Useful), (Quality Work-Quality), (Ease-Difficult), (Access-Quality), and (Instruction-Decision) would improve model fit. These controls make conceptual sense because these pairs of question items included the same word or phrase and are located next to each other in the questionnaire. For instance, question #5 is 'The AAISD Benchmark Assessment Reports are of high quality' associated with the perceived information quality. Question #6 is 'The AAISD Benchmark Assessment Reports help to improve the quality of my work' associated with the perceived usefulness. Even though these two questions were asking about two different factors, respondents were likely to focus on the common word of 'quality'. In addition, these two questions were adjacent to each other so it would be possible that order effects resulted in correlated responses.

The Wald (W) statistic was also used an index for model trimming. This statistic showed that the model (including 6 additional paths among error terms) fit would also not improve by dropping any parameter.

Table 7. Summary of Model-Fit Statistics (Measurement Models)

Model	X^2	df	p-value	CFI	SRMR	RMSEA	90% C.I. RMSEA
Initial CFA	366.44	197	<.001	.940	.048	.065	.054, .075
Final CFA	288.08	191	<.001	.965	.044	.050	.038, .061

The standardized measurement model including factor loadings and residual paths is included in Table 8. Factor loadings are interpreted as regression coefficients (i.e., factor loadings estimate the direct effects of the factors), while residuals are squared to indicate the variance unexplained. Factor loadings are standardized and all are significant ($p<.05$).

Table 8. Standardized Factor Loadings, R^2, Error Paths in the Final Measurement Model (* = p<.05)

Factors and Indicators	Factor Loadings	R^2	Error Path
Perceived Usefulness (F1)			
Student Analysis (V1)	.839*	.704	.544
Quality Work (V2)	.897*	.804	.443
Instruction (V3)	.895*	.801	.446
Decision (V4)	.871*	.759	.491
Useful (V5)	.922*	.849	.388
Perceived Ease of Use (F2)			
Ease (V6)	.730*	.533	.683
Difficult (V7)	.589*	.347	.808
Understand (V8)	.773*	.598	.634
Attitude (F3)			
Fun (V9)	.643*	.413	.766
Dislike (V10)	.812*	.660	.583
Like (V11)	.901*	.812	.434
Intention to Use (F4)			
Continue (V12)	.817*	.667	.577
Increase (V13)	.791*	.626	.612
Problem Solving (V14)	.798*	.636	.603
Share (V15)	.638*	.407	.770
Perceive Information Quality (F5)			
Access (V16)	.641*	.410	.768
Amount (V17)	.797*	.635	.604
Quality (V18)	.755*	.570	.656
External Pressure (F6)			
Administrator (V19)	.855*	.731	.518
Central Office (V20)	.653*	.426	.757

Test of the Structure Model

Once the measurement model was determined to measure the adequacy of all latent constructs, the hypothesized structural model was tested. The structural model sought to determine the factors involved and the extent to which the factors were causally related to the teachers' intention to use student performance information. The hypothesized model included regression effects of one variable or latent factor on another as well as on all six latent factors from the measurement model, along with 3 measures indicating teacher experience, similar prior experience, and user-guide rating. The hypothesized structural model is illustrated in Figure 6.

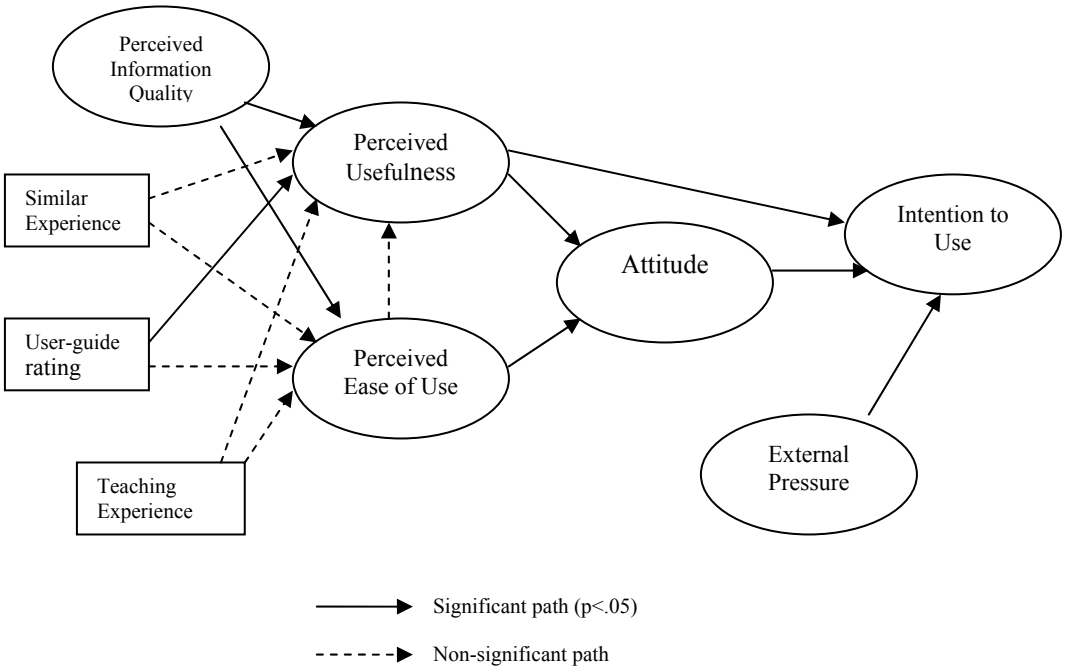


Figure 5. Hypothesized Structural Model.

The hypothesized structural model did not include any direct or indirect effects of two individual differences (teaching experience and prior similar experience) on either perceived usefulness and perceived ease of use. This would seem to contradict earlier studies (Gick & Holyoak, 1987; Harrison & Rainer, 1992; Thompson et al., 1994) yet seem to confirm a recent study finding (Agawal & Prasad, 1999) which also found that there was no significant relationship between experience variables and perceived usefulness and ease of use.

Furthermore, some fit indices for the hypothesized structural model were not within acceptable ranges. For example, The CFI for the model was .920, where it was lower than .960, a reasonable cut-off for model fit (Hu & Bentler, 1999). Another fit indicator that has gained acceptance for structural equation models, was higher than the acceptable limit (RMSEA<.09). In addition, we want a non-significant decrease in model fit by applying the theoretical structural model. The difference of X^2 from the final measurement model to the hypothesized model showed that there was significant decrease in fit (X^2_{diff}= 155.375, df_{diff}=28, p<.001).

By repeating the Lagrange Multiplier test and Wald test for adding other covariance and trimming some of them, the modified structural model emerged without the relationship of teaching experience and prior similar experience toward perceived usefulness and perceived ease of use. Instead, the modified

model added covariances between the user-guide rating variable and perceived information quality, and a direct path from the user-guide rating variable to intention to use. Whereas prior studies (Bostrom, Olfman, & Sein, 1990; Davis & Bostrom, 1993; Igbaria, Gamers, & Davis, 1995) showed the positive influence (direct effects) of training on usefulness and ease of use beliefs due to reducing uncertainty about information system by providing the features of the information system by training, this study does posit the proxy of training, user-guide rating influence on perceived ease of use through perceived information quality. With regard to the direct path from user-guide rating to intention to use, user-guide rating variable may affect teachers' intention to use student performance information like perceived usefulness because the question item about user-guide rating consists of asking about *usefulness* of the user guide. Table 9 showed the modified structural model improved the goodness of model fit indexes.

Table 9. Summary of Model-Fit Statistics (Structural Models)

Model	X^2	df	p-value	CFI	SRMR	RMSEA	90% C.I. RMSEA
Hypothesized Model	443.455	219	p<.001	.920	.097	.071	.061, .080
Modified Model	349.866	216	p<.001	.952	.081	.055	.044, .065

Findings from Hypotheses

Factor loadings were standardized within the results from the modified structural model (Figure 6), along with their measures of statistical significance.

With these statistics and a confirmed goodness of fit for the modified structural model, the following hypotheses were able to be tested.

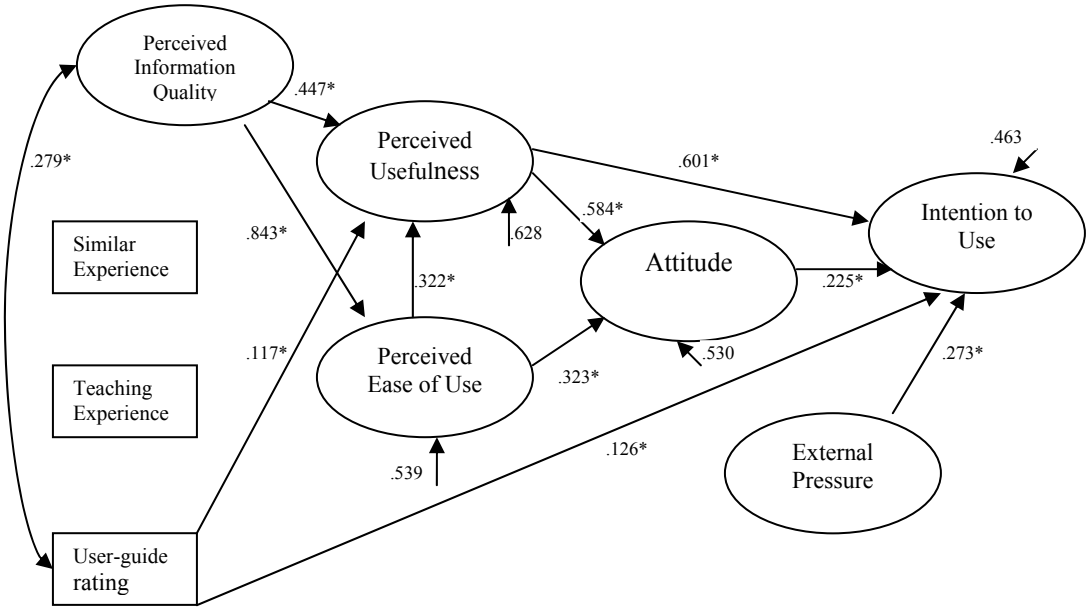


Figure 6. Standardized Modified Structural Model.

H1: Perceived usefulness and ease of use will mediate the influence of selected individual difference and perceived information quality variables on attitude and intention to use student performance information.

An examination of the path coefficients for individual differences shown in Figure 6 reveals that perceived ease of use, perceived information quality, and user-guide rating had a significant positive effect on perceived usefulness, and together, explained 60.6% of the variance in the dependent variable (perceived usefulness). For perceived ease of use, perceived information quality as a

significant determinant and indirect effect of user-guide rating accounted for 71% of the variance in perceived ease of use.

Individual differences in teaching experience and prior similar experience showed no significant direct effect on any other factors. Only user-guide rating was positively associated with perceived usefulness (direct effect), perceived ease of use (spurious relationship), perceived information quality (direct effect), and intention to use (direct effect). Thus, H1 received partial support.

H2: Attitude will mediate the influence of perceived usefulness and ease of use on the intention to use student performance information.

As expected, attitude and perceived usefulness were significant predictors of intentions. In addition, the user-guide rating variable also positively impacted teachers' intention to use student performance information. These three components explained 78.5% of the variance in intention to use.

H3: Perceived ease of use will have a direct effect on attitude and an indirect effect on attitude through perceived usefulness.

Attitude was determined jointly by perceived usefulness and perceived ease of use; these two factors explained 72% of the variance in attitude. As posited in Technology Acceptance Model, perceived ease of use has been a significant predictor of perceived usefulness. However, the results revealed that

there was a non-significant path from perceived ease of use to perceived usefulness in the initially hypothesized model. In the modified model, perceived ease of use was a significant predictor of perceived usefulness (B = .322, p<.05).

H4: Perceived usefulness will have a direct effect on intention to use and an indirect effect on intention through attitude.

Results indicated a statistically significant positive relationship between perceived usefulness and teachers' intention to use student performance information (B = .601, p<.05); between perceived usefulness and attitude (B = .584, p<.05); and between attitude and intention to use (B = .225, p<.05). Path analysis indicated that perceived usefulness had a total effect of .73 on intention to use (Sum of the direct effect and indirect effect through attitude).

Table 10. Direct and Indirect Effects of Usefulness and Ease of Use on Teachers' Intention to Use Student Performance Information

Path	Effect
Usefulness to Intention	
Direct effect	.60
Indirect effect through Attitude	.13
Total effect	**.73**
Ease of Use to Intention	
Indirect effect through Usefulness	.19
Indirect effect through Usefulness and Attitude	.04
Indirect effect through Attitude	.07
Total effect	**.30**

H5: External pressure will have a direct effect on intention to use student performance information.

Results indicated a statistically significant positive relationship between external pressure from school administrators and the central office and teachers' intention to use student performance information (B = .273, $p<.05$)

CHAPTER FIVE

CONCLUSION

SUMMARY

Problem and the purpose of the study

The promotion of use of data by the state government and school districts has changed into a more intense focus on individual school data and new assessment measures. School districts, at the same time, have promoted more and more the use of student performance data through professional development, school improvement planning, and additional incentives for data use. However, there are many barriers to the process of promoting the use of student performance data. What seems to be problematic is that this kind of intensification in the use of data also includes the pressure of the accountability system toward educators.

This research investigated (1) what type of student performance information is available to teachers; (2) how they use the information; and (3) what factors impact positively and/or negatively on teachers' use of student performance information. This study also examined (4) the relationships among perception of usefulness, ease of use, attitude, intention to use student performance information, and external pressure to use the information.

Review of theoretical framework

The theoretical framework for the present research includes 1) consideration of current trends and features of accountability and standards-based school reform movement; 2) exploration of general research streams, such as, Information Diffusion Model, Theory of Planned Behavior, Technology Acceptance Model (TAM), and Information Technology (IT) adoption; 3) discussion of TAM in more depth (Davis, 1989).

TAM, adapted from the Theory of Reasoned Action (Fishbein & Ajzen, 1975), posits that use acceptance is determined by perceived usefulness and perceived ease of use. Perceived usefulness is defined as the extent to which a person believes that using a particular technology (performance information) will enhance job performance (teaching and learning), while perceived ease of use is defined as the degree to which a person believes that using a technology will be free from effort. This study attempts to test the impacts of two basic components in TAM.

Major questions and hypotheses

The major hypothesis was that perceived usefulness, perceived ease of use, attitude, external pressure, perceived information quality, and individual

differences in teaching experiences and training experiences would significantly affect teachers' intention to use student performance information.

- H1: Perceived usefulness and ease of use will mediate the influence of selected individual difference and perceived information quality variables on attitude and intention to use student performance information.

- H2: Attitude will mediate the influence of perceived usefulness and ease of use on intention to use student performance information.

- H3: Perceived ease of use will have a direct effect on attitude and an indirect effect on attitude through perceived usefulness.

- H4: Perceived usefulness will have a direct effect on intention to use and an indirect effect on intention through attitude.

- H5: External pressure will have a direct effect on intention to use student performance information.

Procedures

This study employed both survey research and individual focused interview methods. The perceptions of teachers using student performance information were investigated through a survey of 206 teachers from 16 campuses in an urban school district in Texas. The survey also captured qualitative data from open-ended questions. Individual focused interviews were conducted with one teacher and one staff member from central office to help interpret and explain the quantitative findings. Data analysis included structural equation modeling.

Findings

Findings from Research Question #1 regarding type of student performance information were: The school district responded in an organized way to rising standards by improving alignment within their local curriculum while ensuring that local standards were consistent with those of the state and state mandated tests. Schools and classrooms also were engaged in ongoing test administration to improve student achievement under their own circumstances.

With regard to Research Question #2, there were 4 major themes in teachers' responses on the use of student performance information. Teachers used the benchmark assessment information for (1) checking the efficacy of local curriculum and instructional practices; (2) assessing state curriculum standards and/or objective mastery, (3) preparing for state mandated tests, and (4) deciding NOT to use the benchmark assessment reports, or to use them at the minimum level as directed by the school district.

With regard to Research Question #3, the quality of the test and reports and timely feedback (Information Access) were two common dominant categories that emerged for this study. Teachers also reported other categories as positive or negative factors to impact on their use of benchmark assessment reports (see Table 11).

Table 11. Factors to Impact on Teachers' Use of Benhcmark Assessment Reports

Positive Factors	Negative Factors
Identification of student needs	Poor quality of test and reports
Alignment of curriculum and tests	User unfriendly format
Preparation for instruction	Low information access
Information access,	Time consumption
Information quality	External pressure

Major findings from research hypotheses (Research Question #4) on the Technology Acceptance Model were: (1) the teachers' perception of intention to use student performance information was affected mainly by perceived usefulness, which was conceptually aligned with the first three positive factors in Table 10. (2) Perceived information quality and external pressure impacted directly and/or indirectly on teachers' intention to use benchmark assessment reports. (3) Only the user-guide rating variable in individual difference category showed a statistically significant relationship with perceived usefulness, perceived ease of use, intention to use, and perceived information quality.

CONCLUSION

The success of student accountability systems rests on the extent to which educators use student performance data to make instructional, curricular, and/or evaluation decisions. To this end, this study employed the Technology

Acceptance Model (TAM) to examine factors that influence teachers' intention to use student performance information.

This study found that TAM is a useful conceptual tool for understanding the relationship among facilitators and barriers associated with teachers' intention to use student performance information.

Teachers were likely to use the benchmark assessment data if they perceived the data would be useful in identifying student needs, aligning curriculum and tests, and preparing for their instruction. Moreover, the perception about the quality of the assessment data in term of complexity and accessibility impacted indirectly on teachers' intention to use the benchmark assessment information. Teachers' intention to use the information was also affected directly by the pressure from central office and school administrators as well as the training.

As states develop more intensified accountability testing systems, school districts are likely to respond by increasing test administration.

While the school district developed a variety of assessment systems, the district office tried to quickly process the results and feed them back to teachers in time to identify and provide extra help for students who need it. However, for many teachers, complying with the requirement of state and district mandated tests in concert (i.e., TAKS-Benchmark, SDAA-Benchmark for special education

kids, state and district reading tests) would increase the amount of time devoted to testing and decrease the time available for instruction under a condition of "test score pollution" (Urdan & Paris, 1994, p.139).

Overwhelming teachers with test administration and analyses of the test results leads to differences in the extent to which teachers use student performance information.

One group of teachers responded they used the benchmark assessment information checking the efficacy of local curriculum and instructional practices, assessing state curriculum standards and/or objective mastery, and preparing for state mandated tests. The other group remarked that they decided NOT to use the benchmark assessment reports, or to use them at the minimum level as directed by the school district because classroom teachers often report a low value and very limited use of the results in their day-to-day classroom instruction.

Limitations of Interpretation

A reminder is offered to the reader pertaining to the limitations of the conclusions from this study. The survey method used in this study allowed participants to ultimately select themselves. There can admittedly be some bias associated with this self-selection process.

The statistics of constructs in the research model were self reported, rather than observed. The self-report nature of this study may shake confidence in its conclusions as self-report data can overestimate impact because teacher can genuinely believe that they are using student performance information while not actually doing so.

In addition, the use of a single district, a small sample size that is below the recommended level, and some non-response bias with regard to teaching experiences do not allow for broad generalizability. The limitation of not meeting the minimum sample size suggestions could be related to the instability of the statistical results. According to tests for non-response bias, the respondents in the sample had more teaching experience than teachers in the school district at large.

While all manifest variables for latent factors displayed excellent psychometric properties, a few factors lost half of their original manifest variables. While this aided in parsimony for the construct and overall model, one cannot ultimately be certain that it was measuring the same construct.

DISCUSSION

A specific research objective guiding the study presented here was to shed further light on the implementation of using student performance information by teachers. Several findings from quantitative and qualitative analyses were

discussed within the concepts of 'capacity-building' and 'internal accountability' retrieved from the literature review in Chapter 2.

Lack of Relevance

Tests are used to provide feedback to individual student (formative) or to make decisions about grade, promotion and graduation (summative). Current state mandated testing systems under performance accountability have shown that summative orientation brings pressure to educators. The school district investigated in this study implemented a benchmark assessment system to provide data back to the campuses or to the staff so that they could provide the appropriate remediation necessary for their students.

As Fullan (2001) stated, however, the benchmark assessment system accompanying the state-mandated test left a question mark on the implementation level. For example, the findings of this study show that irrelevance of the benchmark assessment and untimely feedback of student assessment negatively affected the teachers' intention to use the student assessment information. The issue of irrelevance of the test came out of the difference between the timing of benchmark assessment and the instruction. In other words, if teachers have not been teaching district curriculum in that order, it is going to be completely irrelevant to them when their students take that test.

> Every year when we hit fractions, we have to teach fractions for about two weeks. And then, kind of back up and do all over again just because fraction is one of those concepts that's very difficult. So we almost have to go back in, repeat a lot of what we've done with fractions. We don't have to do with every unit by any means. But when we hit that point, we kind of had stand still, so then when we get to benchmark, there is the stop between fractions and benchmark that maybe our kids had not covered yet. So, it's kind of irrelevant.

Untimely Feedback

The school district (AAISD-pseudonym) generated benchmark assessment reports including teachers' individual reports, classroom reports, campus reports, vertical team (Learning Community) reports, and district reports. The reports gave demographic information as well as testing results presented in an Excel spreadsheet. AAISD distributed the benchmark assessment reports to teachers through its internal email system rather shipping hard copies because of time and cost issues. The results of state mandated test, TAKS, were also provided as a CD format so that campuses could get a variety of reports on the TAKS test. Central administrator reasoned that it was cost prohibitive to provide hard paper copies to district personnel. He said, "We calculated the paper. It got to be a half million pages of reports."

Nonetheless, a significant number of teachers who participated in the survey reported concerns about the late return of the test data, which resulted in low intention to use the benchmark assessment reports. In addition to the

technical problems, the findings implied management problems in the process of getting timely feedback.

Some teachers responded that they did not receive data for individual student on the benchmark. Despite the fact that there was a district expectation that all teachers receive individual feedback in an Excel spreadsheet, this did not happen on all campuses.

> If the principal has not chosen to share with how to get in there, then yes that's blocking. Now the principal may want to provide the reason, say I am going to let you get in there, I am going to let you get in there, but I am not going to let you, you, you, No.

Assessment costs school time, effort, and money. AAISD identified that an electronic delivery system was of value to them and discontinued a conventional way to deliver student performance information via hard copies. AAISD tried to balance the depth of the assessment system, richness of data, cost, and utility of the assessment system. For example, AAISD chose a benchmark assessment system developed by one of the departments within AAISD, not a commercial testing system because of the cost and the district's need to control the benchmark assessment system.

Little Allocation Time for Collaboration

Teachers frequently reported having to juggle multiple tasks and to balance competing demands on their time. One of the greatest concerns in using the benchmark assessment reports under pressure is lack of time.

> We need more time to look at benchmark data, it's also we need more time to do curriculum mapping, and we need more time to do wide initiatives, and … All the stuff that we need more time for so I don't know if there is going to be a real solution to that as long as our district's wanting to do so much at one time, then I don't think we are able to get the time to develop one thing to do right. It's always going to be that we are spread so thinly over these other things. And you just have to decide which thing is not going to be that important today. And that's the thing you to let go.

Rigorous state mandated testing has resulted in local school districts producing similar types of assessments which are administered multiple times. As a result, current standard-based school reforms tend to practice a "more is better" approach to assessment. Of concern is that they are using too much assessment without allocating time for reflection and collaboration. If the testing calendar of AAISD (see Appendix J) could be seen, it would be hard to implement Cizek & Rachor (1994) planned assessment system, "one in which each assessment activity that occurs is conducted for well a articulated purpose, clearly defined benefits of assessment, real, tangible, and valuable information, and administration in an efficient manner" (Cizek, 1995, p.247). What seems to be

problematic, even in a well designed model, is that this kind of intensification of data use always includes the pressure of accountability system toward educators.

Alternative Approach: Account-able

From the viewpoint of teachers, the right to choose what they teach, the way to instruct students, and feedback showing whether students know what they have been taught will lead to a greater capacity for improving student achievement. At a more basic level, focusing on those teachers' viewpoints "leads inherently to self-examination- that is, basing one's methods of instruction on internal factors- and relying less on external measures in assessing accountability, such as public inspection (Glenn, 2001, p. 26)."

The literature review suggested the need to supplement or expand effective use of student performance information for developing commitment among teachers for internal accountability in at least three areas. First, focusing on multiple indicators of success in terms of 'density' and 'complexity' might help teachers move beyond test scores. Wheelock (2000) suggests that to improve teaching and learning, accountability systems should focus not only on test scores but on classroom practices that lead to success for students. Another way to supplement the accountability system is to support the notion that the collaboration among teachers increases a sense of mutual support and

responsibility. This notion might lead to the promotion of teachers' commitment to (a) a reflective dialogue on the results of student achievement between teachers and campus administrators, (b) the assistance of campus support teams, and (c) the development of training pertinent to the concerns and issues from their own consciousness (Shelor, 2000). Rallis and MacMullen (2000) also provide similar recommendations defining six activities as the "inquiry cycle (p.770)." (1) Participants discuss and define the outcomes for which they accept accountability. (2) Participants identify important questions concerning student learning. (3) Participants locate the data that exits because of external mandates or routine recording and organize them for their own use. (4) Participants conduct mindful analyses of the data in light of the desired outcomes, and interpret information in light of the school's purposes. (5) Participants take action based on the following questions: What is actually happening? What practice should we continue, and how can we strengthen them. (6) Participants assess the effects of actions. Then, the inquiry cycle begins anew. A third avenue is to set a cornerstone for the essential school improvement questions, such as "Are we skilled enough to use assessment either to keep all learners from losing confidence in themselves to begin with or to rebuild that confidence once it has been destroyed? (Stiggins, 1999, p.196)" This avenue comes through continuous classroom assessment so

that important decisions about student learning are based on specific information about how individual students are performing.

In conclusion, the major practical challenge of current accountability and standards-based school reform is to identify and expand the role of educators in an active and positive way, such as an approach that insures educators are "*account-able*", rather than merely "*held accountable*." This approach may help teachers and principals engage in collaborative inquiry and action for the success of school improvement. As a result of their commitment to the process, account-able educators will engage in formative evaluation, assessment-curriculum alignment, and politically sophisticated campus improvement planning for students' learning.

IMPLICATIONS FOR RESEARCH

This research explored whether the extended Technology Acceptance Model (TAM) was a useful conceptual tool for understanding the relationships among variables associated with teachers' intention to use student performance information. After modifying the core TAM, which was mainly focused on the teacher's perception of usefulness, ease of use, attitude, and intention to use of student performance information, this research extended the core TAM by adding the external pressure, perceived information quality, and three individual

characteristics which might explain some of the individual teachers' intentions to use the test reports within a mandatory adoption situation.

The statistical analyses pointed to the validity of the relationship between individual differences (only the user-guide rating variable) and beliefs as framed in the TAM. They also further validated the remainder of TAM's relationships between beliefs, attitude, and behavioral intentions. The failure of disconfirmation of these relationships has several important implications.

First, a substantial amount of variance (78.5%) in teachers' intention to use student performance information was accounted for by perceived usefulness, perceived ease of use, user-guide rating, perceived information quality, and external pressure. The research model included social factors similar to subjective norm, and facilitating conditions such information quality, resulting in increase of explanation power of teachers' intention factor.

Second, as TAM proposed, both perceived usefulness and perceived ease were important in the acceptance of student performance information by teachers. However, their relative importance in the acceptance process has been shown to be different in prior work. For instance, Davis (1993) found that usefulness dominated ease of use, whereas Adams, Nelson, and Todd (1992) found ease of use to be more influential than usefulness. Davis et al. (1989) found that perceived ease of use was not a significant determinant of intentions immediately

after subjects were exposed to a new technology, but was significant for the same subjects and technology 14 weeks later. In this study, perceived usefulness dominated ease of use (see Table 10). A possible explanation of the dominance of perceived usefulness might be that the benchmark assessment reports were distributed for the first time in 2002-2003, which is still considered as a beginning stage of implementing the benchmark assessment reports.

Third, only the user-guide rating variable of the three individual difference variables had significant effects on perceived usefulness, perceived ease of use, intention to use, and perceived information quality. Insofar as greater tenure in the workplace is a surrogate for age, the researcher expected number of years in teaching to be negatively associated with perceived ease of use student performance information. An alternative explanation might be that age is not a significant influence on ease of use beliefs.

The prior, similar experiences of using local school district email system, productive software and personal computer in general, had no effects on both usefulness and ease of use. One of the reasons might be that there was little variation in the responses of three question items. Most teachers responded that they used the email system, productive software, and computer often in general so that the prior, similar experiences have no effect at all.

The lack of direct effect of user-guide rating variable on perceived ease of use might derive from the following explanation. The survey asked about the degree of 'usefulness' of benchmark workbook (user guide) as a proxy of training because the school district did not provide any training session for using benchmark assessment reports to teachers (see Appendix I). This proxy variable might lead teachers to consider the concept of perceived usefulness of the benchmark assessment reports.

Recommendation for Further Research

From the perspective of theory development, this study found support for the TAM as an adequate and parsimonious conceptualization of acceptance behavior and salience of usefulness and ease of use. The results relating to the weak relationship between individual differences and two core constructs, usefulness and ease of use, suggested that further research should construct simpler models that exclude individual differences altogether. Instead, further research could include variables or factors relating to capacity-building for the implementation of student performance information, such as time allocation.

Another recommendation follows when the results with regard to the relative importance of those two core constructs, perceived usefulness and ease of use, are compared to others. Future research could be focused on identifying

reasons to believe that the relative strength of the two constructs is a consequence of the stage of implementing benchmark assessment system, which was administered for the first time in 2002-2003. For instance, Davis et al. (1989) found that perceived ease of use was not a significant determinant of intentions immediately after subjects were exposed to a new technology, but was significant for the same subjects and technology 14 weeks later. Future research is needed to help clarify the relative importance of perceived usefulness and ease of use of student performance information at the different phase of implementing benchmark assessment system.

In addition, the benchmark assessments are administered more than twice a year. Therefore, there is a need to study the TAM in the longitudinal way, following up the change of external pressure under intensifying student performance accountability.

IMPLICATIONS FOR PRACTICE

School districts should provide professional development on how to use benchmark assessment reports to the majority of classroom teachers. While the school district investigated in this study is implementing the benchmark assessment system as a response to the new state-mandated testing program, much of this training targeted only school administrators. School administrators then

transferred what they had learned to a selected group of teachers such as those on school improvement teams. If the majority of classroom teachers were simply receiving direction from others in a limited way, such as a user guide on the website, it would be doubtful that it would help teachers to promote use of student performance information.

School administrators should allocate more time or modify existing schedules so that teachers may analyze and reflect on student performance data, plan revisions to their curriculum and teaching practices, and receive professional development on how to use the information effectively.

School districts and school administrators should develop an attitude in assessment, "less is more," by limiting the quantity of student assessments in order to accomplish those two practical implications mentioned earlier. The attitude of "less is more" means a kind of affordability that teachers can typically have enough time to examine student assessment results and apply their reflection on the results to their instructional practices. This affordability may accompany with a method of action research in increasing resource, time allocation for reflection on student performance information. The term action research refers to teams of teachers studying a situation and identifying a problem, considering a range of possible solutions to the problem, experimenting with one or more solutions, and assessing the results of the experimentations.

In addition, “less is more” attitude could apply to the quality of the benchmark assessment reports. The reports generated by the school district created an often overwhelming amount of information to which teachers and administrators are expected to respond. Therefore, school districts need to reduce the amount of the benchmark assessment information, leaving teachers to gauge student progress based on their internal assessments, such as teacher made tests and quizzes.

The findings relating to external pressure that teacher perceived mainly emerged from the fear of district and school administration’s evaluating teachers’ performance by comparing test results and the anxiety on their students’ failure in promotion. School districts and campus administration should continue to explore strategies for reducing the external pressure.

Overall, the key to school improvement using student performance information is to provide teachers with more resources, expertise, training, and support, in a word, “capacity.” This approach could let teachers view accountability as a helpful tool that seeks to improve schooling by using of student performance information as a diagnostic devise, increasing collaboration across campuses and classrooms, and making more effective use of school resources. Teachers should be usually people who make things happen, NOT people to whom things happen.

APPENDICES

Appendix A: Approval Letter

Department of Assessment and Research
[redacted]
[redacted], Texas [redacted]

Phone: [redacted]

December 12, 2002

Mr. Sungkwan Yang
3378 Lake Austin Blvd. #E
Austin, Texas 78703

Dear Mr. Yang:

Please be advised that your request to conduct a research study in the [redacted] Independent School District ([redacted]ISD) is approved. The title of your research is "Teacher's perception on student performance information". The [redacted] Independent School District has placed great importance on dissemination of data and data analysis. Your research topic is of very high interest to the [redacted] Independent School District.

This letter of research approval will serve to introduce you to each of our campuses and principals as you proceed with your activities. The Assessment and Research Office continues to guard and safeguard the valuable limited time that our teachers and principals have to participate in research activities, therefore, we ask that you be sensitive to this issue.

We hope that you will be able to conduct your research successfully and we wish you the very best in your efforts.

Sincerely,

[redacted]
Director of Assessment and Research

Appendix B: Invitation Letter to Participate in Survey

Dear Faculty Members;

I am a doctoral student at the University of Texas at Austin and a graduate assistant for the Educational Productivity Council. The following survey is a part of a research project being conducted to fulfill the requirements of a Ph.D. dissertation at the University of Texas at Austin. I hope to learn whether teachers' intention to use the AAISD Benchmark Assessment Reports is affected by teachers' perception about usefulness and ease of use of AAISD Benchmark Assessment Reports and organizational factors.

You have been randomly selected. You will be one of 680 teachers who participate in this survey.

Participation in this survey is voluntary and on an anonymous and confidential basis. You may decide not to complete the survey or may decide not to answer specific questions, but your attempt to answer each question is very much appreciated by the researcher. All information obtained in connection with this survey will remain confidential and anonymous.

Your decision whether or not to participate will not prejudice your future relations with The University of Texas at Austin or your school district. You may keep a copy of this cover letter. Responding to the survey indicates a willingness to participate in the study.

If you have any questions about this study, I will be happy to answer them. I can be contacted through The University of Texas at Austin, Sanchez Building (SZB) Rm. 360A, (512) 475-8597; my home phone, (512) 478-7919; or electronic mail, ysksh@mail.utexas.edu.

Thank you for your participation.

Sung-Kwan Yang
Principal Investigator

Appendix C: Consent Form to Participate in Interview

Teachers' Perception of Use of Student Performance Information

You are invited to participate in my dissertation study about teachers' use of student performance information. I am a doctoral student at The University of Texas at Austin, Department of Educational Administration, and this study supports my dissertation. I hope to learn whether teachers' use of student performance information is affected by teachers' perception about student performance information and organizational factors. You are randomly selected.

If you decide to participate, your will be interviewed in depth about your experiences while using student performance information. The interview will be audio taped. Taped interview will require approximately a half hour.

There will be no risks to you if you participate. All information obtained in connection with this interview will remain confidential and your individual comments will be disclosed only with your permission and anonymously. Identifiers on audio tapes will be recorded in a manner that will not reveal your identity. Additionally, audio tapes will not be released to anyone and will be destroyed upon completion and approval of this dissertation.

Your decision whether or not to participate will not prejudice your future relations with The University of Texas at Austin and your school district. If you decide to participate, you are free to discontinue participation at any time during the study without prejudice. Simply notify me of your decision.

If you have any questions about this study or this consent form, please ask me. If you have any additional questions during or after the study, I will be happy to answer them. I can be contacted through The University of Texas at Austin, Sanchez Building (SZB) Rm. 360A, (512) 475-8597; my home phone, (512) 478-7919; or electronic mail, ysksh@mail.utexas.edu.

You are making a decision whether or not to participate. Your signature indicates that you have read the information provided above and have decided to participate. If you later decide to discontinue participation in this study after signing this form, you may withdraw at any time without prejudice. You may keep a copy of this form.

Participant Signature	Printed Name	Phone Number	Date

Sung-Kwan Yang, Researcher	Jay D. Scribner, Supervising Professor
SZB 360A (see phone number above)	SZB 310 (471-7551)

Appendix D: Questionnaire

Teacher Survey

Perception of Use of
The AAISD Benchmark Assessment Reports

This is a survey of teachers' perception of use of the AA Independent School District (AAISD) Benchmark Assessment Reports. This is **not** a survey of *the Benchmark Tests*. Please read the following statements and then respond by placing a check mark or X in the box that best matches your opinion on the issue addressed in the statement. Please provide only one response to each statement.

The AAISD Benchmark Assessment Reports	Very Strongly Disagree	Moderatelly Disagree	Slightlly Disagree	Neutral	Slightlly Agree	Moderatelly Agree	Very Strongly Agree
1. Are easily accessible.	①	②	③	④	⑤	⑥	⑦
2. Contain an appropriate amount of information.	①	②	③	④	⑤	⑥	⑦
3. Are easy to use.	①	②	③	④	⑤	⑥	⑦
4. Help me to accomplish student analysis.	①	②	③	④	⑤	⑥	⑦
5. Are of high quality.	①	②	③	④	⑤	⑥	⑦
6. Help to improve the quality of my work.	①	②	③	④	⑤	⑥	⑦
7. Are *difficult* to use.	①	②	③	④	⑤	⑥	⑦
8. Improve my instruction of students.	①	②	③	④	⑤	⑥	⑦
9. Make it easy for me to make instructional decisions.	①	②	③	④	⑤	⑥	⑦
10. Are useful.	①	②	③	④	⑤	⑥	⑦
11. Are understandable	①	②	③	④	⑤	⑥	⑦

		Very Strongly Disagree	Moderatelly Disagree	Slightlly Disagree	Neutral	Slightlly Agree	Moderatelly Agree	Very Strongly Agree
12.	I intend to continue using the AAISD Benchmark Assessment Reports.	①	②	③	④	⑤	⑥	⑦
13.	My fellow teachers believe that I should use the AAISD Benchmark Assessment Reports.	①	②	③	④	⑤	⑥	⑦
14.	The AAISD Benchmark Assessment Reports are fun to use.	①	②	③	④	⑤	⑥	⑦
15.	My use of AAISD Benchmark Assessment Reports is voluntary.	①	②	③	④	⑤	⑥	⑦
16.	I intend to increase my use of the AAISD Benchmark Assessment Reports in the future.	①	②	③	④	⑤	⑥	⑦
17.	My campus administrators believe that I should use the AAISD Benchmark Assessment Reports.	①	②	③	④	⑤	⑥	⑦
18.	I intend to use the information from the AAISD Benchmark Assessment Reports for problem solving and decision making.	①	②	③	④	⑤	⑥	⑦
19.	I *dislike* using the AAISD Benchmark Assessment Reports.	①	②	③	④	⑤	⑥	⑦
20.	I intend to share data from the AAISD Benchmark Assessment Reports with other teachers.	①	②	③	④	⑤	⑥	⑦
21.	Central office personnel believe that I should use the AAISD Benchmark Assessment Reports.	①	②	③	④	⑤	⑥	⑦
22.	I like using the AAISD Benchmark Assessment Reports.	①	②	③	④	⑤	⑥	⑦
23.	The AAISD Benchmark Assessment Reports provide an attractive working environment.	①	②	③	④	⑤	⑥	⑦

Please answer the following questions about yourself by checking the space in front of the appropriate information or providing the information requested.

1. Male_____ Female _____
2. Please specify your age range. A.___21-30 B.___31-40 C.___41-50 D.___51 or over
3. How many years of teaching experience do you have? (years)
4. How many years of teaching experience do you have in the AA ISD? (years)
5. What grade level(s) do you teach? ( )
6. What subject(s) do you teach?
   A.__ English B.__ History C.__ Language Arts D.__ Math E.__ Science
   F.__ Social Studies G.__ Elementary (All subjects) H.__ preschool

7. Is **the Benchmark Workbook** useful to use AAISD Benchmark Assessment Reports
   A.__ Yes B.__ No C.___ Never see it

8. I currently use the Personal Computer for **Lotus Notes**.

1	2	3	4	5	6	7
☐	☐	☐	☐	☐	☐	☐
Never Used One			*Used One Occasionally*			*Used One A Lot*

9. I currently use the Personal Computer for **Productive Software**.
   This includes: Word processing programs (e.g. MS Word, Word Perfect ... etc), Spreadsheet programs (e.g. MS Excel, Lotus 123 ... etc), Presentation programs (PowerPoint ... etc) and others.

1	2	3	4	5	6	7
☐	☐	☐	☐	☐	☐	☐
Never Used One			*Used One Occasionally*			*Used One A Lot*

10. Overall, I would rate my experience with the Personal Computer as

Please write in the information.

List the **positive** factors to impact on your use of the AAISD Benchmark Assessment Reports
_____
_____
_____

List the **negative** factors to impact on your use of the AAISD Benchmark Assessment Reports
_____
_____
_____

How do you actually use AAISD Benchmark Assessment Reports?
_____
_____
_____

Thank you for taking the time to complete this questionnaire.

Appendix E: Interview Protocol

Interview Protocol

Participant Information

1. Course subject
2. Level of campus
3. Similar prior experience
4. Years of teaching
5. Participation in training

Interview Guide

1. What types of assessment does your school currently use?
    a. Subjectivity
    b. Method
    c. Purposes
    d. Frequency about using the assessment results

2. How do you actually use the student performance information?
    a. Alignment among test results, curriculum, and instruction
    b. Contribution of student assessment information to campus improvement planning process.

3. What positive and negative factors influence your use of student performance information?
    a. Priority of learning about the assessment data and using it in their school and/or district
    b. Any formal or informal mechanisms in place at their campus to support the use of student assessment data, including professional development opportunities to support their use of student assessment information

Appendix F: Correlation Matrix for test of Full Model

Correlation Matrix

	PU1	PU2	PU3	PU4	PU5
PU1	1.0000				
PU2	.7766	1.0000			
PU3	.7609	.8313	1.0000		
PU4	.7206	.7703	.8379	1.0000	
PU5	.7547	.7368	.7997	.8209	1.0000
PEU1	.5539	.4031	.3419	.4146	.4659
PEU2	.4442	.3588	.3018	.3341	.3794
PEU3	.5939	.4738	.4596	.5331	.6480
AT1	.4787	.4869	.4552	.4386	.5041
AT2	.5484	.5881	.5440	.5395	.6461
AT3	.6316	.6392	.6176	.6152	.7069
IU1	.5942	.5618	.5937	.6115	.6848
IU2	.5313	.5597	.5983	.5731	.6311
IU3	.6047	.6234	.6708	.6192	.6586
IU4	.4647	.4351	.4834	.4737	.5006
Q1	.3857	.3831	.4223	.3930	.4101
Q2	.5704	.4893	.5095	.4943	.5514
Q3	.6391	.6149	.5076	.5217	.5874
EP3	.2824	.2889	.2818	.2998	.2823
EP4	.1048	.0845	.1162	.1765	.1292

	PEU1	PEU2	PEU3	AT1	AT2
PEU1	1.0000				
PEU2	.5983	1.0000			
PEU3	.5671	.4506	1.0000		
AT1	.3942	.3336	.3717	1.0000	
AT2	.4692	.4352	.5190	.4782	1.0000
AT3	.4549	.3715	.5607	.5991	.7327
IU1	.3562	.2538	.4997	.3513	.5030
IU2	.3296	.2569	.4117	.4066	.4551
IU3	.2960	.2417	.4066	.3805	.5547
IU4	.3152	.2175	.3612	.3317	.4054
Q1	.4105	.3509	.3993	.2511	.3151

Q2	.5727	.4109	.5290	.2914	.4576
Q3	.4594	.3739	.4925	.4517	.4606
EP3	.2059	.0808	.2360	.1048	.1815
EP4	.1850	-.0450	.2422	-.0181	.0870

	AT3	IU1	IU2	IU3	IU4
AT3	1.0000				
IU1	.5244	1.0000			
IU2	.5128	.6745	1.0000		
IU3	.5488	.6259	.6383	1.0000	
IU4	.5052	.5114	.5116	.4956	1.0000
Q1	.3751	.2755	.2940	.2561	.2606
Q2	.4592	.4052	.3802	.3730	.2606
Q3	.5032	.4750	.5188	.4996	.3048
EP3	.2017	.4191	.2923	.3032	.3288
EP4	.0378	.2764	.1539	.2043	.2016

	Q1	Q2	Q3	EP3	EP4
Q1	1.0000				
Q2	.5558	1.0000			
Q3	.3423	.5627	1.0000		
EP3	.1847	.1538	.1758	1.0000	
EP4	.0872	.1013	.0527	.5584	1.0000

Appendix G: Descriptive Statistics

Descriptive Statistics

	N	Mean	Std.	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
Useful1	206	4.38	1.723	-.413	.169	-.973	.337
Useful2	206	3.89	1.665	-.205	.169	-1.015	.337
Useful3	206	4.09	1.596	-.333	.169	-.820	.337
Useful4	206	4.11	1.652	-.303	.169	-.865	.337
Useful5	206	4.19	1.668	-.490	.169	-.807	.337
Ease1	206	4.39	1.591	-.297	.169	-.799	.337
Ease2	206	4.18	1.618	.054	.169	-.757	.337
Ease3	206	4.70	1.398	-.583	.169	-.030	.337
Attitude1	206	2.39	1.342	.402	.169	-1.136	.337
Attitude2	206	3.80	1.666	.077	.169	-.558	.337
Attitude3	206	3.58	1.546	-.012	.169	-.710	.337
Intention1	206	4.91	1.429	-.646	.169	.331	.337
Intention2	206	4.16	1.574	-.306	.169	-.272	.337
Intention3	206	4.70	1.509	-.741	.169	.218	.337
Intention4	206	4.60	1.607	-.555	.169	-.211	.337
Info Quality1	206	4.81	1.617	-.645	.169	-.496	.337
Info Quality2	206	4.74	1.545	-.539	.169	-.622	.337
Info Quality3	206	4.13	1.648	-.316	.169	-.876	.337
Pressure3	206	5.90	1.223	-.965	.169	.183	.337
Pressure4	206	6.00	1.286	-1.274	.169	1.014	.337
Experience	206	12.558	8.4075	.616	.169	-.438	.337
Exp in RRISD	206	7.476	6.6124	1.463	.169	1.988	.337
Valid N (listwise	206						

Middle School

Campus Summary - Grade 6
Writing - Fall
Test Administered: October 2002

6th Grade - Writing

Average # and % Correct		
# Items	# Correct	% Correct
20	14	68%

% Mastery by Objective			
3	4	5	6
67%	13%	44%	49%

% at Passing and Proficient Standard
17% Proficient
57% Passing

NOTE: Passing = 70% and above;
Proficient - 85% and above

							Written Composition			
							1	2	3	4
Teacher	COURSE	GRADE	SECTION	Student ID	Last Name	First Name	Idea Dev.	Org.	Voice	Conventions
120	6712	6	3				3	2	3	0
120	6712	6	3				3	3	2	0
120	6712	6	3				1	2	1	0
120	6712	6	3				3	2	2	0
120	6712	6	3				2	1	2	0
120	6712	6	3							
120	6712	6	3							

							Mastery of TAKS Objective				Correct Items by TAKS Objective			
							4/5	3/4	5/6	4/5				
Teacher	COURSE	GRADE	SECTION	Student ID	Last Name	First Name	3	4	5	6	3	4	5	6
120	6712	6	1				Y	N	N	Y	5	2	3	4
120	6712	6	1				Y	N	Y	Y	5	2	5	4
120	6712	6	1				Y	N	N	N	4	3	4	2
120	6712	6	1				Y	N	N	N	5	3	3	3
120	6712	6	1				Y	N	N	Y	4	3	4	4
120	6712	6	1											

Teacher	COURSE	GRADE	SECTION	Student ID	Last Name	First Name	Eth Code	SP ED	EOD	Total Items	Total Correct	% Correct	PASS	PROF
120	6712	6	3				2	N		20	15	75%	Y	N
120	6712	6	3				5	N		20	15	75%	Y	N
120	6712	6	3				5	N		20	16	80%	Y	N
120	6712	6	3											

Middle School
6th Grade Writing - Item Analysis
Fall Benchmark
Test Administered: October 2002

Item #	TAKS Obj.	TEKS	Correct Answer	% Responses			
				1	2	3	4
1	6	7.16B	3	1%	1%	81%	17%
2	6	7.16B	1	77%	0%	2%	21%
3	3	7.18C	2	1%	73%	13%	14%
4	6	7.18H	4	14%	26%	10%	51%
5	6	7.17G	3	6%	3%	46%	44%
6	3	7.18C	2	8%	78%	9%	5%
7	4	7.17B	3	13%	31%	53%	3%
8	5	7.17C	2	1%	74%	22%	3%
9	5	7.17F	3	8%	1%	90%	1%
10	3	7.18C	4	8%	4%	2%	87%
11	4	7.17A	1	63%	12%	3%	23%
12	5	7.17C	2	6%	64%	2%	28%
13	4	7.17E	4	12%	12%	12%	64%
14	3	7.18D	3	21%	2%	73%	4%
15	4	7.17A	4	5%	17%	36%	42%
16	6	7.16C	1	88%	4%	0%	8%
17	5	7.17D	3	3%	2%	90%	5%
18	5	7.17C	3	1%	4%	69%	26%
19	3	7.18C	1	72%	6%	12%	9%
20	5	7.17H	1	30%	3%	5%	61%

Middle School
6th Grade Writing - Item Analysis
Fall Benchmark - Teacher 120
Test Administered: October 2002

Item #	TAKS Obj.	TEKS	Correct Answer	% Responses			
				1	2	3	4
1	6	7.16B	3	0%	0%	78%	22%
2	6	7.16B	1	76%	0%	0%	24%
3	3	7.18C	2	0%	78%	11%	11%
4	6	7.18H	4	19%	30%	8%	43%
5	6	7.17G	3	3%	0%	59%	38%
6	3	7.18C	2	0%	97%	3%	0%
7	4	7.17B	3	5%	16%	78%	0%
8	5	7.17C	2	0%	76%	24%	0%
9	5	7.17F	3	3%	0%	97%	0%
10	3	7.18C	4	3%	3%	0%	95%
11	4	7.17A	1	62%	11%	0%	27%
12	5	7.17C	2	3%	65%	0%	32%
13	4	7.17E	4	3%	8%	14%	76%
14	3	7.18D	3	14%	0%	84%	3%
15	4	7.17A	4	3%	16%	43%	38%
16	6	7.16C	1	92%	0%	0%	8%
17	5	7.17D	3	0%	0%	97%	3%
18	5	7.17C	3	0%	0%	73%	27%
19	3	7.18C	1	76%	5%	8%	11%
20	5	7.17H	1	38%	5%	3%	54%

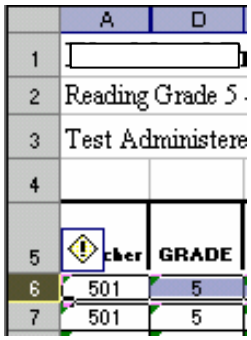
Appendix I: User Guide-Benchmark Workbook

About the Benchmark Workbook

After opening the data from Lotus Notes be sure to save it to your harddrive, network server or a disk.

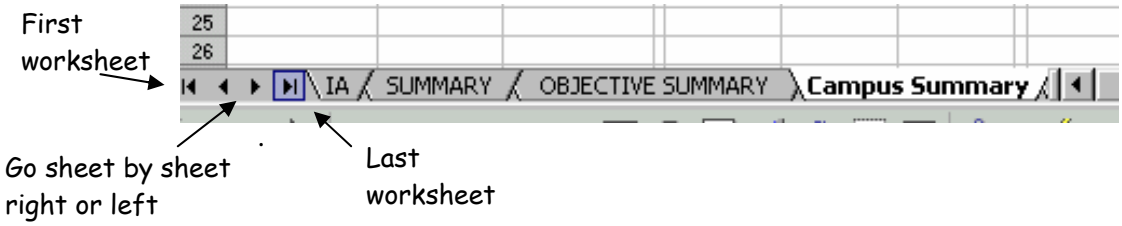
Basics

Cells in a spreadsheet are identified by column and row. There will be times when you want to select an entire row or column. By clicking on 6 in the gray strip, you are selecting the entire row. You can select a column by clicking on the letter.



Navigation

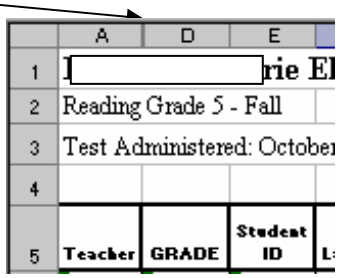
When you first open the Excel document, you are looking at the campus summary for a given grade level. At the bottom of the screen are tabs for the various worksheets. On the left, at the bottom are arrows which will allow you to move to other tabs.



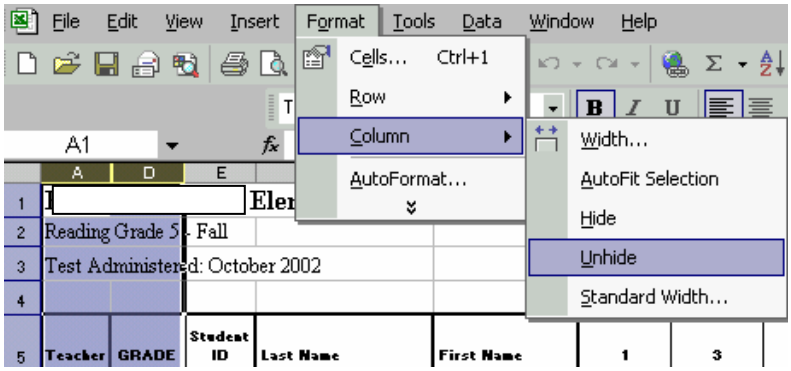
You select a different worksheet by clicking on that tab. After selecting a worksheet, you can use the horizontal or vertical scroll bars to navigate within the worksheet. In addition, you can click in a cell and use arrow keys to move through the worksheet.

Hidden columns

Due to the extensive information in the document, some columns are hidden from view. For example, in this screenshot, columns B & C are hidden.



To show the hidden columns, click on A, hold the button down and drag to D, so both are highlighted. Then go to **Format/Column/Unhide**, see below.

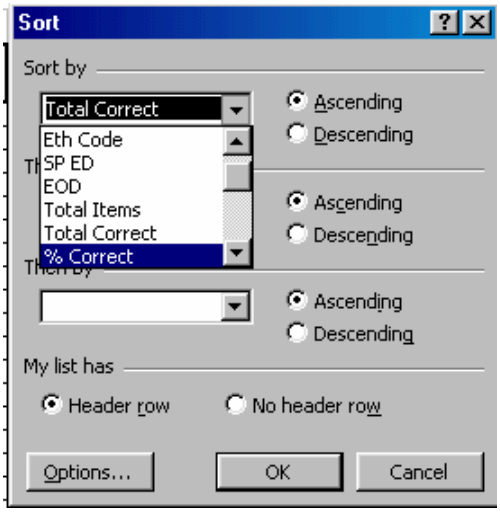


Sorting

One of the advantages of working with a spreadsheet is the ability to sort the information in various ways. **Before sorting on a particular worksheet, make sure hidden columns are unhidden for that worksheet.** (see Hidden Columns above) The first screen shot shows a section of a Summary worksheet. (Student names and personal information are not shown due to confidentiality issues.) Initially, the information is listed by student last name. As a teacher, I would like to see the information sorted by % Correct. Starting with the first student in my class, I will click on the row numbers that appear on the far left of the screen, dragging down through the last student. By clicking on the row numbers, you are selecting all the information for that row. (Note: if you are not the first class, see Additional Tips on how to split the window so you can see the column titles.) Once the information is highlighted, go to Data/Sort. A window will be displayed allowing you to select the method by which you will sort the data.

Before

Eth Code	SP ED	EOD	Total Items	Total Correct	% Correct
4	N	N	21	15	71%
3	N	Y	21	19	90%
5	N	N	21	21	100%
5	N	N	21	20	95%
5	N	N	21	17	81%
5	N	N	21	15	71%
5	N	N	21	20	95%
5	N	N	21	17	81%
3	N	N	21	15	71%
3	N	N	21	17	81%
3	N	N	21	8	38%
3	N	Y	21	17	81%
5	Y	N	21	20	95%
5	N	N	21	19	90%
5	N	N	21	19	90%
5	N	N	21	17	81%
5	N	N	21	18	86%
4	N	N	21	19	90%
5	N	N	21	18	86%
5	N	N	21	10	48%
3	N	N	21	15	71%



In this example, I've used the drop-down to highlight (select) % Correct. The radio button to the right indicates that the data will be in Ascending order. (lowest to highest) If you prefer Descending order, simply click on the radio button next to that. Then click on OK

After

Eth Code	SP ED	EOD	Total Items	Total Correct	% Correct
5	N	N	21	8	38%
5	N	N	21	10	48%
5	N	N	21	15	71%
5	N	N	21	15	71%
3	N	Y	21	15	71%
4	N	N	21	15	71%
3	N	N	21	17	81%
5	N	N	21	17	81%
5	N	N	21	17	81%
4	N	N	21	17	81%
5	N	N	21	17	81%
5	N	N	21	18	86%
5	N	N	21	18	86%
5	Y	N	21	19	90%
5	N	N	21	19	90%
3	N	Y	21	19	90%
5	N	N	21	19	90%
3	N	N	21	20	95%
3	N	N	21	20	95%
5	N	N	21	20	95%
3	N	N	21	21	100%

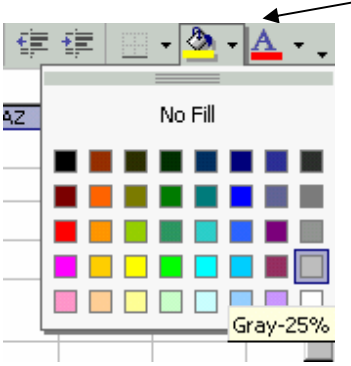
As you can see, the % Correct column now reflects the new sort. At this point, you may want to insert blank rows, or colored rows as dividers between certain areas. Directions on the next page will show you how to do that.

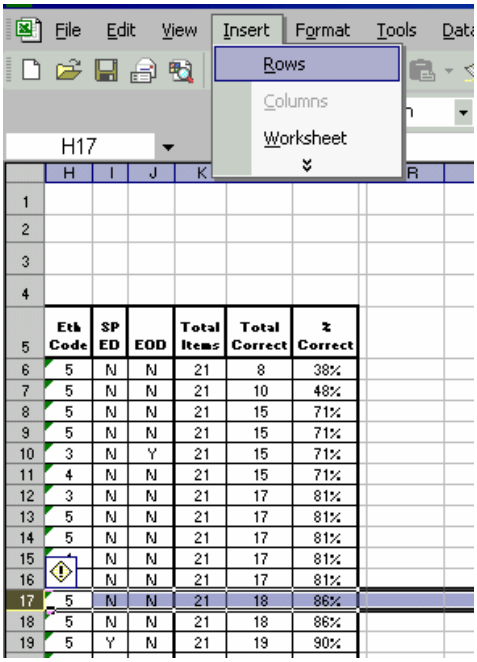
Adding a row

I want to identify the breaks between certain % Correct ranges. Specifically I want to differentiate between Passing, 70% and above and Proficient 85% and above. To do that I will insert a Row and color it. The row will need to be between rows 16 and 17.

	Eth Code	SP ED	EOD	Total Items	Total Correct	% Correct
5						
6	5	N	N	21	8	38%
7	5	N	N	21	10	48%
8	5	N	N	21	15	71%
9	5	N	N	21	15	71%
10	3	N	Y	21	15	71%
11	4	N	N	21	15	71%
12	3	N	N	21	17	81%
13	5	N	N	21	17	81%
14	5	N	N	21	17	81%
15	4	N	N	21	17	81%
16	5	N	N	21	17	81%
17	5	N	N	21	18	86%
18	5	N	N	21	18	86%
19	5	Y	N	21	19	90%

The row I select will be shifted down and the inserted row will take its place. I will select row 17 by clicking on 17. That highlights the entire row. Then go to Insert/Rows.
A blank row will appear between 16 and 18.
While 17 is still highlighted, go up to the paint bucket on the toolbar and select a color.





The screen shot to the right shows how the worksheet looks after a row has been added. I would also add a row between rows 7 & 8, because that would show those not passing.

	Eth Code	SP ED	EOD	Total Items	Total Correct	% Correct
5						
6	5	N	N	21	8	38%
7	5	N	N	21	10	48%
8	5	N	N	21	15	71%
9	5	N	N	21	15	71%
10	3	N	Y	21	15	71%
11	4	N	N	21	15	71%
12	3	N	N	21	17	81%
13	5	N	N	21	17	81%
14	5	N	N	21	17	81%
15	4	N	N	21	17	81%
16	5	N	N	21	17	81%
17						
18	5	N	N	21	18	86%
19	5	N	N	21	18	86%
20	5	Y	N	21	19	90%

Suggested Ways to Sort

Objective Summary –

Sort by Objective: In this document you can sort to show student mastery by objective.

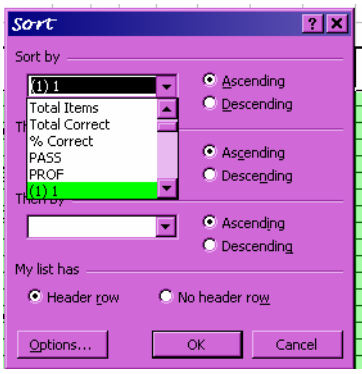
*Note: Unhide all columns before attempting to sort. (See Hidden columns on page 1)

Before sorting.

In looking at Objective 1, I want to sort the information so all students not mastering the objective will be grouped together. Select students by clicking on the row and dragging down to highlight all your students. In the menu bar, go to Data/Sort. The Sort window will come up. Click on the drop-down arrow to select the column being used for sorting. The column is next to the Prof column so I'll select (1)1.

Mastery of TAKS Objective			Correct Items by TAKS Objective		
4/5	5/7	5/7			
1	3	4	1	3	4
Y	N	Y	4	3	7
Y	Y	Y	5	6	6
Y	Y	Y	5	7	7
Y	Y	Y	5	6	7
Y	Y	Y	5	6	5
N	Y	Y	2	6	6
Y	Y	Y	5	6	7
Y	Y	Y	4	5	6
Y	Y	Y	5	5	5
Y	Y	Y	4	6	7
N	N	N	1	4	3
Y	Y	N	5	7	3
Y	Y	Y	5	7	7
Y	Y	Y	4	7	6
Y	Y	Y	5	6	6
Y	Y	Y	4	7	6
Y	Y	Y	4	7	6
Y	Y	Y	5	6	6
Y	Y	Y	4	6	6
N	N	N	3	3	3
N	Y	Y	2	6	6
Y	Y	N	4	7	4

Note: Sometimes the column heading will be a letter instead of words.



	Mastery of TAKS Objective		
	4/5	5/7	5/7
PROF	1	3	4
0	Y	N	Y
1	Y	Y	Y
1	Y	Y	Y
1	Y	Y	Y
0	Y	Y	Y
0	N	Y	Y
1	Y	Y	Y
0	Y	Y	Y
0	Y	Y	Y
0	Y	Y	Y
0	N	N	N
0	Y	Y	N
1	Y	Y	Y
1	Y	Y	Y
1	Y	Y	Y
0	Y	Y	Y
1	Y	Y	Y

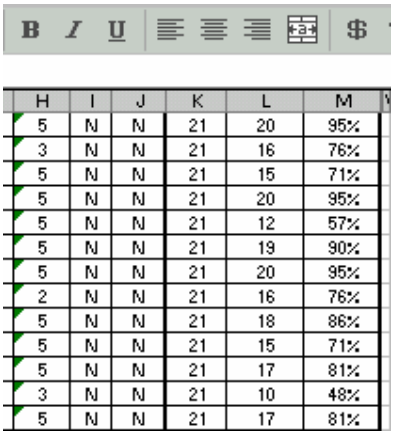
Once the information has been sorted, the students who have not mastered objective 1 are listed first. In this example, the first four students did not master objective 1. You can follow the same procedure for Objectives 3 and 4.

Mastery of TAKS Objective			Correct Items by TAK	
4/5	5/7	5/7		
1	3	4	1	3
N	Y	Y	2	6
N	N	N	1	4
N	N	N	3	3
N	Y	Y	2	6
Y	N	Y	4	3
Y	Y	Y	5	6
Y	Y	Y	5	7
Y	Y	Y	5	6
Y	Y	Y	5	6
Y	Y	Y	5	6
Y	Y	Y	4	5
Y	Y	Y	5	5
Y	Y	Y	4	6
Y	Y	N	5	7
Y	Y	Y	5	7
Y	Y	Y	4	7
Y	Y	Y	5	6
Y	Y	Y	4	7
Y	Y	Y	4	7
Y	Y	Y	5	6
Y	Y	Y	4	6
Y	Y	N	4	7

Additional Tips

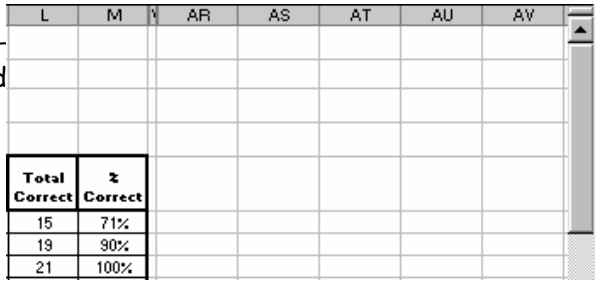
Due to the size of spreadsheets, it is challenging at times to determine what information goes with which column. Here's a handy trick.

In the shot to the right, I have scrolled down the page and column headers are not visible. There is a way to "split the page" so that headers can be seen.



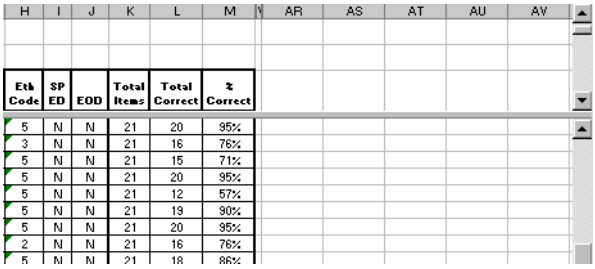
H	I	J	K	L	M
5	N	N	21	20	95%
3	N	N	21	16	76%
5	N	N	21	15	71%
5	N	N	21	20	95%
5	N	N	21	12	57%
5	N	N	21	19	90%
5	N	N	21	20	95%
2	N	N	21	16	76%
5	N	N	21	18	86%
5	N	N	21	15	71%
5	N	N	21	17	81%
3	N	N	21	10	48%
5	N	N	21	17	81%

By pointing the mouse cursor at the spot indicated to the right, you'll get a two-headed arrow.
Once there is a two-headed arrow, the user can click, hold the mouse button down, and drag down.



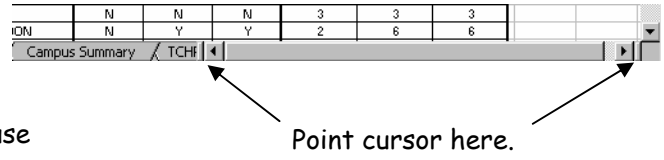
L	M	AR	AS	AT	AU	AV
Total Correct	% Correct					
15	71%					
19	90%					
21	100%					

The shot to the right shows two vertical scroll areas. The upper area is set to show the column headers. In the lower area, the user can scroll to the bottom and still know what information is in a given column.



H	I	J	K	L	M	AR	AS	AT	AU	AV
Eth Code	SP ED	EOD	Total Items	Total Correct	% Correct					
5	N	N	21	20	95%					
3	N	N	21	16	76%					
5	N	N	21	15	71%					
5	N	N	21	20	95%					
5	N	N	21	12	57%					
5	N	N	21	19	90%					
5	N	N	21	20	95%					
2	N	N	21	16	76%					

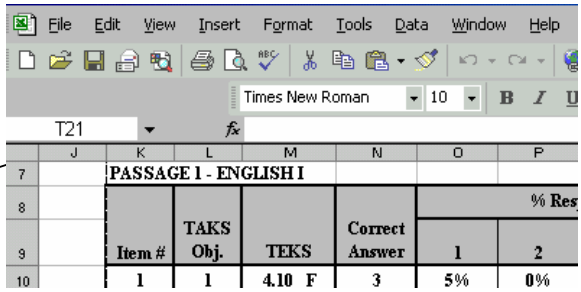
Just as a screen can be split vertically, it is also possible to do so horizontally. The shot to the right shows the horizontal scroll bar. Arrows indicate where to point the mouse cursor.



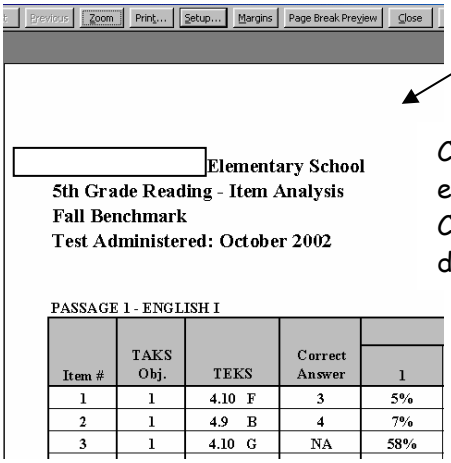
Printing

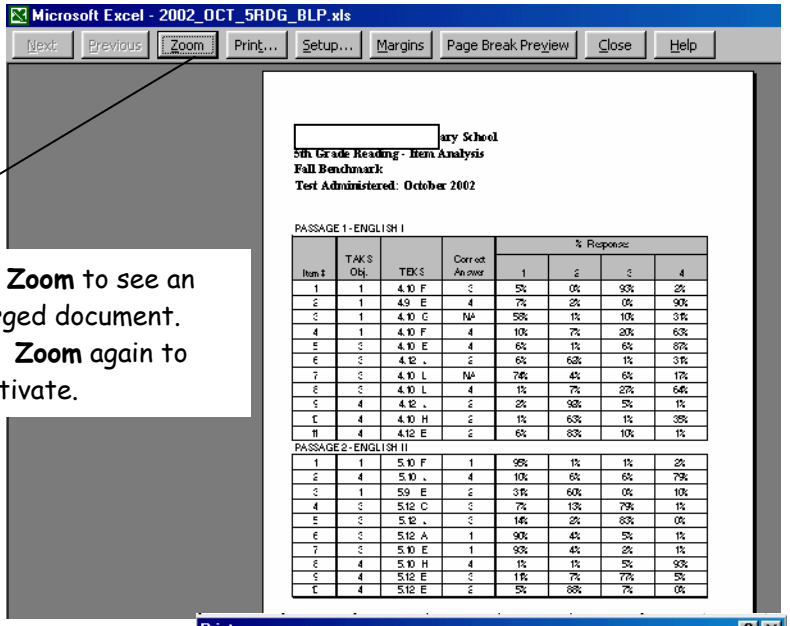
If the user wants to see how a worksheet will appear when printed, click on the Print Preview icon located on the toolbar.

Print Preview



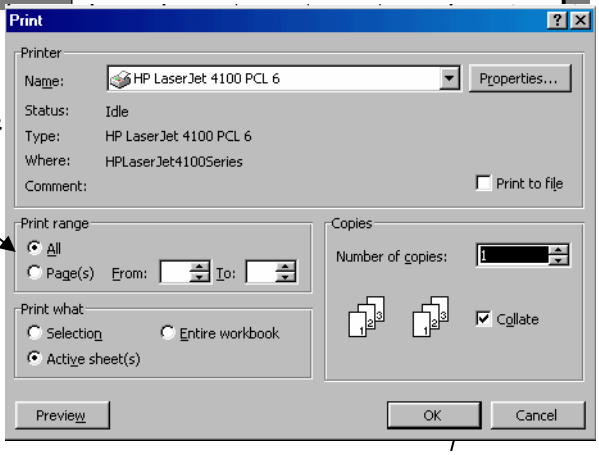
The shot to the right is how this worksheet will print.





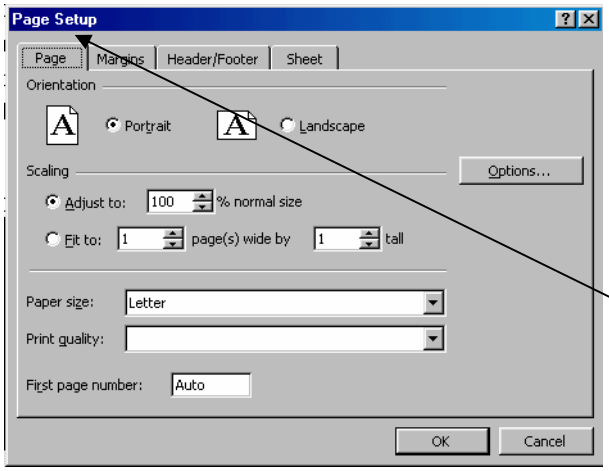
Click **Zoom** to see an enlarged document. Click **Zoom** again to deactivate.

The **Print** button brings up the print window, allowing the user to select which pages will be printed and the number of copies to be printed.



Print range. Print all or select pages.

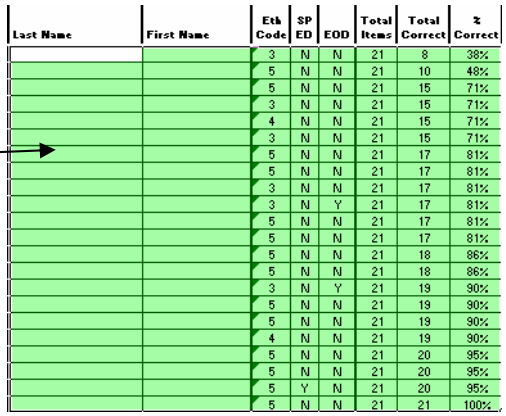
Number of copies



Setup brings up the window to the left allowing for changes in pages, margins, and so forth.

Setting Print Area

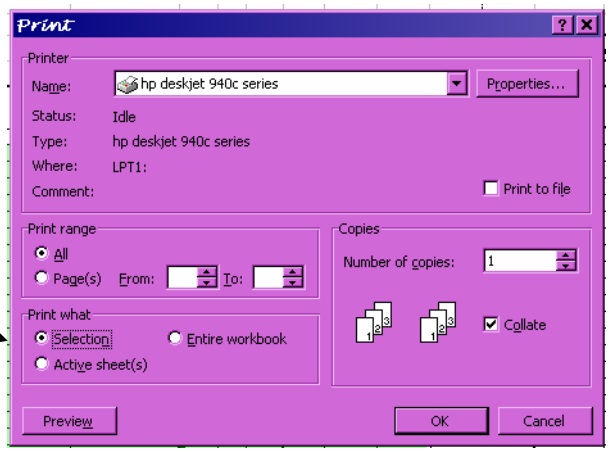
It is possible to select and print a specific area of a worksheet. Simply highlight the area you wish to print.



Last Name	First Name	Eth Code	SP ED	EOD	Total Items	Total Correct	% Correct
		3	N	N	21	8	38%
		5	N	N	21	10	48%
		5	N	N	21	15	71%
		3	N	N	21	15	71%
		4	N	N	21	15	71%
		3	N	N	21	15	71%
		5	N	N	21	17	81%
		5	N	N	21	17	81%
		3	N	N	21	17	81%
		3	N	Y	21	17	81%
		5	N	N	21	17	81%
		5	N	N	21	17	81%
		5	N	N	21	18	86%
		5	N	N	21	18	86%
		3	N	Y	21	19	90%
		5	N	N	21	19	90%
		5	N	N	21	19	90%
		4	N	N	21	19	90%
		5	N	N	21	20	95%
		5	N	N	21	20	95%
		5	Y	N	21	20	95%
		5	N	N	21	21	100%

Go to **File/Print**.

At the Print screen, click the radio button next to Selection in the Print What area.



Only the area you have selected will be printed.

Appendix J: District Testing Calendar

Due to the important nature of color to show the difference of test schedules the following pages are consecutively numbered as:

2003-04 Elementary Benchmark Calendar

Some dates shown are testing "windows" and do not reflect the specific date on which a test is given at individual campuses.

July 2003

S	M	T	W	Th	F	S
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30	31		

August 2003

S	M	T	W	Th	F	S
					1	2
3	4	5	6	7	8	9
10	11	12	13	14	15	16
17	18	19	20	21	22	23
24	25	26	27	28	29	30
31						

September 2003

S	M	T	W	Th	F	S
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30				

October 2003

S	M	T	W	Th	F	S
			1	2	3	4
5	6	7	8	9	10	11
12	13	14	15	16	17	18
19	20	21	22	23	24	25
26	27	28	29	30	31	

November 2003

S	M	T	W	Th	F	S
						1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30						

December 2003

S	M	T	W	Th	F	S
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30	31			

READING (Key: Orange box)

Nov 3 - 7	Gr. 3, 4, 5
Jan 12 - 16	Gr. 3
Mar 1 - 5	Gr. 4 & 5
Apr 28	Gr. 2

WRITING (Key: Highlighted green)

Oct 20 - 24	Gr. 3-5
Nov 3 - 7	Gr. 2
Dec 1 - 5	Gr. K & 1
Jan 12 - 16	Gr. 4
Feb 24	Gr. 2, 3, 5
Mar 1 - 5	Gr. K & 1
Apr 28	Gr. 2

SCIENCE (Key: Pink box)

Dec 1 - 18	Gr. K-5
Apr 29 - May 14	Gr. K-5

MATH (Key: Underlined blue)

Oct 13 - 24	Gr. K-5
Mar 1 - 12	Gr. K-5
Apr 27	Gr. 2

SOCIAL STUDIES (Key: Purple text)

Oct 7 - 18	Gr. K-3
Oct 13 - 24	Gr. 4 & 5
Jan 7 - 17	Gr. K-3
Feb 24 - Mar 7	Gr. K-3
Mar 22 - Apr 2	Gr. 4 & 5
May 1 - 16	Gr. K-3

January 2004

S	M	T	W	Th	F	S
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	30	31

February 2004

S	M	T	W	Th	F	S
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29						

March 2004

S	M	T	W	Th	F	S
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30	31			

April 2004

S	M	T	W	Th	F	S
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	30	

May 2004

S	M	T	W	Th	F	S
						1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31					

June 2004

S	M	T	W	Th	F	S
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30			

2003-04 Elementary Testing Calendar

Some dates shown are testing "windows" and do not reflect the specific date on which a test is given at individual campuses.

July 2003

S	M	T	W	Th	F	S
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30	31		

August 2003

S	M	T	W	Th	F	S
					1	2
3	4	5	6	7	8	9
10	11	12	13	14	15	16
17	18	19	20	21	22	23
24	25	26	27	28	29	30
31						

September 2003

S	M	T	W	Th	F	S
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30				

,

October 2003

S	M	T	W	Th	F	S
			1	2	3	4
5	6	7	8	9	10	11
12	13	14	15	16	17	18
19	20	21	22	23	24	25
26	27	28	29	30	31	

November 2003

S	M	T	W	Th	F	S
						1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30						

December 2003

S	M	T	W	Th	F	S
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30	31			

FALL STUDY (Key: Highlighted blue)

Sept 15 – 19	Gr. 3-5

DRA (Key: Blue text)

Aug 25 – Sept 19	Gr. 1-5
Jan 6 – 29	Mid-year for students below level in fall
May 1 - 23	End of year

KDRA (Key: Underlined blue)

Jan 6 – 29	K
May 1 - 23	K

SDAA (Key: Orange box)

Feb 24	Gr. 4 writing
Apr 27	Gr. 3-5 math
Apr 28	Gr. 3-5 reading
May 3 – 14	Gr. 4 writing
	Gr. 3-5 reading & math

TAKS (Key: Highlighted yellow)

Jan 22 – Feb 3	Field tests Gr. 4 writing
	Field tests Gr. 4 Spanish writing
Feb 24	Gr. 4 writing, Eng. & Spanish
Mar 3	Gr. 3 reading – Eng. & Spanish
Mar 15 – 26	Spanish field tests
Apr 27	Gr. 3-5 math, Eng. & Spanish
Apr 28	Gr. 3 reading retest – Eng. & Span.
	Gr. 4-5 reading, Eng. & Spanish
Apr 29	Gr. 5 science – Eng. & Span.
June 29	Gr. 3 reading retest – Eng. & Span.
June 21 July 16	Gr. 3 reading alternative assessment, Eng. & Spanish

RPTE (Key: Green box)

Mar 15 - 26	Gr. 3-5

January 2004

S	M	T	W	Th	F	S
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	30	31

February 2004

S	M	T	W	Th	F	S
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29						

March 2004

S	M	T	W	Th	F	S
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30	31			

April 2004

S	M	T	W	Th	F	S
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	30	

May 2004

S	M	T	W	Th	F	S
						1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31					

June 2004

S	M	T	W	Th	F	S
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30			

2003-04 Middle School Benchmark Calendar

Some dates shown are testing "windows" and do not reflect the specific date on which a test is given at individual campuses.

July 2003

S	M	T	W	Th	F	S
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30	31		

August 2003

S	M	T	W	Th	F	S
					1	2
3	4	5	6	7	8	9
10	11	12	13	14	15	16
17	18	19	20	21	22	23
24	25	26	27	28	29	30
31						

September 2003

S	M	T	W	Th	F	S
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30				

October 2003

S	M	T	W	Th	F	S
			1	2	3	4
5	6	7	8	9	10	11
12	13	14	15	16	17	18
19	20	21	22	23	24	25
26	27	28	29	30	31	

November 2003

S	M	T	W	Th	F	S
						1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30						

December 2003

S	M	T	W	Th	F	S
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30	31			

READING (Key: Pink box)

Sep 2 – 12	Gr. 6-8
Feb 2 – 13	Gr. 6-8
Mar 3 – 7	Gr. 6-8

WRITING (Key: Highlighted green)

Aug 25 – Sep 5	Gr. 6-8
Jan 6 – 30	Gr. 6-8

SCIENCE (Key: Orange box)

Dec 1 - 12	Gr. 6-8
Apr 28 - May 14	Gr. 6-8

MATH (Key: Underlined blue)

Nov 10 - 21	Gr. 6-8
Mar 1 - 12	Gr. 6-8

SOCIAL STUDIES (Key: Purple text)

Nov 10 - 14	Gr. 7
Dec 8 - 12	Gr. 6 & 8
Apr 1 - 7	Gr. 7

January 2004

S	M	T	W	Th	F	S
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	30	31

February 2004

S	M	T	W	Th	F	S
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29						

March 2004

S	M	T	W	Th	F	S
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30	31			

April 2004

S	M	T	W	Th	F	S
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	30	

May 2004

S	M	T	W	Th	F	S
						1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31					

June 2004

S	M	T	W	Th	F	S
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30			

2003-04 Middle School Testing Calendar

Some dates shown are testing "windows" and do not reflect the specific date on which a test is given at individual campuses

July 2003

S	M	T	W	Th	F	S
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30	31		

August 2003

S	M	T	W	Th	F	S
					1	2
3	4	5	6	7	8	9
10	11	12	13	14	15	16
17	18	19	20	21	22	23
24	25	26	27	28	29	30
31						

September 2003

S	M	T	W	Th	F	S
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30				

October 2003

S	M	T	W	Th	F	S
			1	2	3	4
5	6	7	8	9	10	11
12	13	14	15	16	17	18
19	20	21	22	23	24	25
26	27	28	29	30	31	

November 2003

S	M	T	W	Th	F	S
						1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30						

December 2003

S	M	T	W	Th	F	S
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30	31			

FALL STUDY (Key: Highlighted blue)

Sept 15 – 19	Gr. 6-8

GATES-MACGINITIE (Key: Underlined blue)

Jan 6 – 30	Gr. 8

SDAA (Key: Orange box)

Sep 15 – 19	Gr. 6-8 fall study
Feb 24	Gr. 7 writing
Apr 27	Gr. 6-8 math
Apr 28	Gr. 6-8 reading
May 3 – 14	Gr. 7 writing
	Gr. 6-8 reading
	Gr. 6-8 math

TAKS (Key: Purple text)

Sep 15 – 19	Gr. 6-8 fall study
Jan 20 – Feb 3	Field tests Gr. 7 writing
Feb 24	Gr. 7 writing
Apr 27	Gr. 6-8 math
	Gr. 6 Span. math
Apr 28	Gr. 6-8 reading
	Gr. 6 Span reading
Apr 29	Gr. 8 social studies

RPTE (Key : Highlighted green)

Mar 15 - 26	Gr. 6-8

ALGEBRA QUALIFYING TEST (Key: Highlighted pink)

Aug 6	Gr. 7 & 8

January 2004

S	M	T	W	Th	F	S
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	30	31

February 2004

S	M	T	W	Th	F	S
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29						

March 2004

S	M	T	W	Th	F	S
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30	31			

April 2004

S	M	T	W	Th	F	S
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	30	

May 2004

S	M	T	W	Th	F	S
						1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31					

June 2004

S	M	T	W	Th	F	S
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30			

2003-04 High School Benchmark Calendar

Some dates shown are testing "windows" and do not reflect the specific date on which a test is given at individual campuses.

July 2003

S	M	T	W	Th	F	S
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30	31		

August 2003

S	M	T	W	Th	F	S
					1	2
3	4	5	6	7	8	9
10	11	12	13	14	15	16
17	18	19	20	21	22	23
24	25	26	27	28	29	30
31						

September 2003

S	M	T	W	Th	F	S
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30				

October 2003

S	M	T	W	Th	F	S
			1	2	3	4
5	6	7	8	9	10	11
12	13	14	15	16	17	18
19	20	21	22	23	24	25
26	27	28	29	30	31	

November 2003

S	M	T	W	Th	F	S
						1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30						

December 2003

S	M	T	W	Th	F	S
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30	31			

ENG LANGUAGE ARTS (Key: Pink box)

Aug 25 – Sep 5	Gr. 9-11
Jan 6 - 30	Gr. 9-11

SCIENCE (Key: Highlighted green)

Dec 1 - 12	Gr. 9-12
Apr 28 - May 14	Gr. 9-12

MATH (Key: Highlighted blue)

Nov 10 - 21	Gr. 9-12
Mar 1 - 12	Gr. 9-12

SOCIAL STUDIES (Key: Purple text)

Oct 20 - 31	Gr. 12
Dec 8 - 12	Gr. 9-11
Apr 1 - 7	Gr. 10-12

January 2004

S	M	T	W	Th	F	S
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	30	31

February 2004

S	M	T	W	Th	F	S
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29						

March 2004

S	M	T	W	Th	F	S
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30	31			

April 2004

S	M	T	W	Th	F	S
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	30	

May 2004

S	M	T	W	Th	F	S
						1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31					

June 2004

S	M	T	W	Th	F	S
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30			

2002-03 High School Testing Calendar

Some dates shown are testing "windows" and do not reflect the specific date on which a test is given at individual campuses.

July 2003

S	M	T	W	Th	F	S
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30	31		

August 2003

S	M	T	W	Th	F	S
					1	2
3	4	5	6	7	8	9
10	11	12	13	14	15	16
17	18	19	20	21	22	23
24	25	26	27	28	29	30
31						

September 2003

S	M	T	W	Th	F	S
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30				

October 2003

S	M	T	W	Th	F	S
			1	2	3	4
5	6	7	8	9	10	11
12	13	14	15	16	17	18
19	20	21	22	23	24	25
26	27	28	29	30	31	

November 2003

S	M	T	W	Th	F	S
						1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30						

December 2003

S	M	T	W	Th	F	S
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30	31			

FALL STUDY (Key: Highlighted blue)

Sept 15 – 19	Gr. 9-12

ACT (Key: Blue box)

Sept 27	College entrance exam
Oct 25	College entrance exam
Dec 13	College entrance exam
Feb 7	College entrance exam
Apr 3	College entrance exam
June 12	College entrance exam

SAT I & SAT II (Key: Red text)

Oct 11	College board test
Nov 1	College board test
Dec 6	College board test
Jan 24	College board test
Mar 27	SAT I only
May 1	College board test
June 5	College board test

PSAT (Key: Purple box)

Oct 18	College board test

AP TESTS (Key: Highlighted pink)

May 5 – 16	Advanced placement tests for various subjects

TAKS (Key: Highlighted yellow)

Jan 20 – Feb 3	Field tests Gr. 9 reading
	Field tests Gr. 10 & 11 Eng. lang. arts
Feb 24	Gr. 9 reading
	Gr. 10 & 11 Eng. lang. arts
Apr 27	Gr. 11 math
	Gr. 10 social studies
Apr 28	Gr. 10 math
	Gr. 11 science
Apr 29	Gr. 11 social studies
	Gr. 9 math
	Gr. 10 science

RPTE (Key: Highlighted green)

Mar 15 - 26	Gr. 9-12

TAAS (Key: Orange box)

Oct 21	Exit level writing retest
Oct 22	Exit level math retest
Oct 23	Exit level reading retest
Feb 24	Exit level writing retest
Feb 25	Exit level math retest
Feb 26	Exit level reading retest
Apr 27	Exit level writing retest
Apr 28	Exit level math retest
Apr 29	Exit level reading retest

January 2004

S	M	T	W	Th	F	S
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	30	31

February 2004

S	M	T	W	Th	F	S
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29						

March 2004

S	M	T	W	Th	F	S
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30	31			

April 2004

S	M	T	W	Th	F	S
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	30	

May 2004

S	M	T	W	Th	F	S
						1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31					

June 2004

S	M	T	W	Th	F	S
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30			

REFERENCES

Adams, D., Nelson, R., & Todd, P. (1992). Perceived usefulness, ease of use, and usage of information technology: A replication. *MIS Quarterly, 16 (2)*, 227-247.

Adams, J. & Kirst, M. (1999). New demands and concepts for educational accountability: Striving for results in an era of excellence. In J. Murphy & K. Louis (Eds.), *Handbook of research on educational administration* (2nd ed., pp.463-489). San Francisco: Jossey-Bass.

Agarwal, R. (2000). Individual Acceptance. In R. Zmud (Ed.), *Framing the domains of IT management: Projecting future… through the past*, (pp.85-104). Cincinnati, Ohio: Pinnaflex Educational Resources, Inc.

Agarwal, R. & Prasad, J. (1997). The role of innovation characteristics and perceived voluntariness in the acceptance of information technologies. *Decision Sciences, 28 (3)*, 557-582.

Agarwal, R. & Prasad, J. (1999). Are individual difference germane to the acceptance of new information technologies? *Decision Science, 30 (2)*, 361-391.

Ajzen, I. (1985). From intentions to action: A theory of planned behavior. In J. Kuhl and J. Beckmann (Eds). *Action Control: From Cognition to Behavior*. New York: Springer-Verlag.

Ajzen, I., and Fishbein, M. (1980). *Understanding attitudes and predicting behavior*. Englewood Cliffs, NJ: Prentice-Hall.

Allington, R., & Walmsley, S. (1995). Redefining and reforming instructional support programs for at-risk students. In R. Allington & S. Walmsley (Eds.), *No quick fix* (pp. 19-44). New York: Teachers College Press.

Anderson, G. (1990). *Fundamentals of educational research*. London: The Falmer Press.

Anderson, J. & Gerbing, D. (1988). Structural equation modeling in practice: A review and recommended two step approach. *Psychological Bulletin, 103 (3)*, 411-423.

Baker, E. (1989). *Can we fairly measure the quality of education?* (CSE Tech. Rep. No. 290, Grant No. OERI-G-86-0003). Los Angeles: University of California Center for Research on Evaluation, Standards, and Student Testing.

Barber, M. & Phillips, V. (2000). Big change questions "Should large-scale assessment be used for accountability?" The fusion of pressure and support. *Journal of Educational Change, 1*, 277-281.

Bentler, P. (1995). *EQS structural equations program manual*. Encino, CA: Multivariate Software, Inc.

Blank, R. (1993). Developing a system of education indicators: Selecting, implementing, and reporting indicators. *Educational Evaluation and Policy Analysis, 11 (2)*, 65-80.

Bollen, K. (1989). *Structural equations with latent variables*. New York, NY: John Wiley & Sons.

Boomsma, A. (1982). The robustness of LISREL against small sample sizes in factor analysis models. In K. Joreskog and H. Wolds (Eds.), *Systems under indirect observation: Causality, structure, prediction* (pp.149-173). Amsterdam: North Holland.

Bostrom, R., Olfman, L., & Sein, M. (1990). The importance of learning style in end-user training. *MIS Quarterly, 14 (1)*, 101-119.

Bracey, G. W. (1999). The ninth Bracey report on the condition of public education. Phi Delta Kappan.. 147-68.

Brancheau, J., & Wetherbe, J. (1990). The adoption of spreadsheet software: Testing innovation diffusion theory in the context of end-user computing. *Information Systems Research 1 (2),* 115-143.

Brown, D. (1993). *The political influence of state-mandated testing reform through the eyes of principals and teachers.* (ERIC Document Reproduction Service No. Ed360 737).

Burstein, L., Oakes, J., & Guiton, G. (1992). Education indicators. In M. Alkin (Ed.), *Encyclopedia of educational research* (6th ed., pp. 409-418). New York: Macmillan.

Cannell, J.J. (1987). Nationally normed elementary achievement testing in America's public school: How all 50 states are above the national average. *Educational Measurement: Issues and Practice, 7(4)*, 12-15.

Chubb, J. & Moe, T. (1990). *Politics markets and America's schools*. Washington, DC: The Brookings Institution.

Chubb, J., & Moe, T. (1988). *What price democracy? Politics, markets and American schools.* Washington, DC: The Brookings Institution.

Cizek, G. (1995). The big picture in assessment and who ought to have it. *Phi Delta Kappan, 77 (3)*, 246-249.

Cizek, G, & Rachor, R. (1994). The real testing bias: The role of values in educational assessment. *NASSP Bulletin, 78 (560)*, 83-93.

Clark, M., Madaus, G., Horn, C., & Ramos, M. (2000). Retrospective on educational testing and assessment in the 20th century. *Journal of Curriculum Studies, 32 (2)*, 159-181.

Cohen, J. (1977). *Statistical power analysis for behavioral sciences*. London: Academic Press.

Compeau, D.R., and Higgins, C.A. (1995). Application of social cognigive theory to training for computer skills. *Information Systems Research, 6(2)*, 118-143.

Creemers, B. & Reetzig, G. (1996). School level conditions affecting the effectiveness of instruction. *School Effectiveness and School Improvement, 7,* 197-228.

Creswell, J. (1998). *Qualitative inquiry and research design*. Thousand Oaks, CA: Sage.

Cromey, Allison & Hanson, Matthew (2000). *An exploratory analysis of school-based student assessment systems.* North Central Regional Educational Laboratory.

Cuban, L. (1988). *The managerial imperative: The practice of leadership in schools*. New York: New York University Press.

Darling-Hammond, L. (1988). Accountability and teacher professionalism. *American Educator, 12 (4)*, 8-13, 38-43.

Darling-Hammond, L. (1997). *The right to learn*. San Francisco: Jossey-Bass.

David, J. (1988). The use of indicators by school districts: Aid or threat to improvement? *Phi Delta Kappan, 69*, 499-503.

Davis, F. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *Management Information Science Quarterly, September*, 319-340.

Davis, F. (1993). User acceptance of information technology: System characteristics, user perceptions and behavioral impacts. *International Journal of Man-Machine Studies, 38*, 475-487.

Davis, F., Bagozzi, R., and Warshaw, P. (1989). User acceptance of computer technology: Comparison of two theoretical models. *Management Science, 35(8),* 982-1003.

Davis, S., & Bostrom, R. (1993). Training end users: An experimental investigation of the roles of the computer interface and training methods. *MIS Quarterly, 17 (1)*, 61-85.

Dillman, D. (2000). *Mail and internet surveys: The tailored design method*. New York, NY: John Wiley & Sons, Inc.

Educational Productivity Council (2002). Retrieved January 15, 2003, from University of Texas-Austin, Educational Productivity Council Web site: http://epc.edb.utexas.edu

Elmore, R. & Rothman, R. (1999). *Testing, teaching, and learning: A guide for states and school districts.* Washington DC: National Academy Press.

Elmore, R. (1996). Getting to scale with good educational practice. *Harvard Education Review, 6 (1)*, 1-26.

Elmore, R., Abelmann, C., & Fuhrman, S. (1996). The new accountability in state education reform: From process to performance. In H. Ladd (Ed.), *Holding school accountable: Performance-based reform in education* (pp.65-98). Washington DC: The Brookings Institution.

Elmore, R., Peterson, P., & McCarthey, S. (1996). *Restructuring in the classroom: Teaching learning and school organization*. San Francisco: Jossey-Bass.

Empson, L. (1999, October 8). The challenge of managing knowledge. *The Financial Times Mastering Series: Mastering Strategy*, 8-10.

English, F. (2000). *Deciding what to teach and test* (Millennium Ed.). Thousand Oaks, CA: Corwin Press.

Fichman, R. and Kemerer, C. (1997). The assimilation of software process innovations; An organizational learning perspective. *Management Science, 43(1)*, 1345-1363.

Firestone, W. & Bader, B. (1992). *Redesigning teaching: Professionalism or bureaucracy?* Albany: State University of New York Press.

Firestone, W., Mayrowetz, D., & Fairman, J. (1998). Performance-based assessment and instructional change: The effects of testing in Maine and Maryland. *Educational Evaluation and Policy Analysis, 20 (2)*, 95-113.

Fishbein, M. & Ajzen, I. (1975). *Belief, attitude, intention and behavior: An introduction to theory and research.* Addison-Wesley, Reading, MA.

Fitz, J. & Gorard, S. (2000). School choice and SES stratification of schools: New findings from England and Wales. A paper presented at AERA 2000, New Orleans, LA.

Fitz, J., & Gorard, S. (2000). School choice and SES stratification of schools: New findings from England and Wales. A paper presented at AERA 2000, New Orleans, LA.

Fleck, M. (1999). *Elementary principals' use of student assessment information*. Doctoral dissertation, Teachers College, Columbia University.

Fowler, F. (1993). *Survey research methods*. Newbury Park, CA: Sage.

Fuhrman, S. (1994). *Evaluating of performance in the United States: Changes in accountability*. New Brunswick, NJ: Rutgers University, Consortium for Policy Research in Education.

Fuhrman, S. (2001). Introduction. In S. Fuhrman (Ed.), *From the capital to the classroom* (pp.1-12). University of Chicago Press.

Fullan, M. (2001). *The new meaning of educational change*. New York: Teachers College Press.

Gall, M., Borg, W., & Gall, J. (1996). *Educational research: An introduction (6th ed.).* White Plains, NY: Longman.

Garn, G. (2001). Moving from bureaucratic to market accountability: The problem of imperfect information. *Educational Administration Quarterly, 37 (4)*, 571-599.

Gefen, D. (2000). Applying TAM to a parallel systems convention strategy. *Journal of Information Technology Theory & Application, 2(2)*, available at http://peffers.ba.ttu.edu//jitta/journal/volume2_2/volume2_2p1.htm.

Gick, M. & Holyoak, K. (1987). The cognitive basis of knowledge transfer. In S. Cormier & J. Hagman (Eds.), *Transfer of learning: Contemporary research and applications* (pp.9-46). San Diego, CA: Academic Press.

Glasman, N. (1994). *Making better decisions about school problems: How administrators use evaluation to find solutions*. Thousand Oaks, CA: Corwin.

Glasnapp, D., Poggio, J., & Miller, D. (1991). Impact of "low stakes" state minimum competency testing program on policy, attitudes, and achievement. In R. Stake (Ed.), *Advances in program evaluation: Vol. 1.*

Effects of mandated assessment on teaching (pp. 101-140). Greenwich, CT: JAI Press Ltd.

Glatthorn, A. (1994). *Developing a quality curriculum*. Alexandria, VA: Association for Supervision and Curriculum Development.

Glenn, C. (2001). School reform: Accountability, data use, and the principal. Doctoral dissertation, Johnson & Wales University.

Goertz, M. (2001). Standard-based accountability: Horse trade or horse whip? In S. Fuhrman (Ed.), *From the capital to the classroom* (pp.39-59). University of Chicago Press.

Goertz, M. E. (2000). Implementing standards-based reform: Challenges for state policy. In *Closing the gap. A Report on the Wingspread 1999 Conference. Beyond the Standards Horse Race: Implementation, Assessment, and Accountability – The Keys to Improving Student Achievement.* Washington, DC: Council for Basic Education.

Gordon, S. & Reese, M. (1997). High stakes testing: Worth the price? *Journal of School Leadership, 7*, 345-368.

Grant, S. (2000). Teachers and tests: Exploring teachers' perceptions of changes in the New York State-mandated testing program. *Education Policy Analysis Archives, 8 (14)*. Available: http://epaa.asu.edu/epaa/v8n14.html.

Grant, S. (2001). An uncertain lever: Exploring the influence of state-level testing on teaching social studies. *Teachers College Record, 102 (3)*, 398-426.

Grissmer, D., and Flanagan, A. (1998). *Exploring rapid achievement gains in North Carolina and Texas.* Washington, DC: National Education Goals Panel.

Guthrie, J. (1990). The evolving political economy of education and implications for educational evaluation. *Educational Review, 42 (2).* 109-131.

Hancock, G. & Mueller, R. (2001). Rethinking construct reliability within latent variable systems. In R. Cudeck, S. Du Toit, & D. Sorbom (Eds.), *Structural Equation Modeling: Present and future* (pp.195-216). Lincolnwood, IL: Scientific Software International, Inc.

Haney, W. (1984). Testing reasoning and reasoning about testing. *Review of Educational Research, 54 (4)*, 507-654.

Haney, W. (2000). The myth of the Texas miracle in education. *Education Policy Analysis Archives* [On-line serial], *8 (41)*. Available: http://epaa.asu.edu/epaa/v8n41/

Hansen, J. (1993). Is educational reform through mandated accountability an oxymoron? *Measurement and Evaluation in Counseling and Development, 26 (1),* 11-21.

Harrison, A. & Rainer, R. (1992). The influence of individual differences on skill in end-user computing. *Journal of Management Information Systems, 9 (1)*, 93-111.

Herman, J. & Abedi, J. (1994). Assessing the effects of standardized testing on schools. *Educational & Psychological Measurement, 54 (2)*, 471-482.

Herman, J. (1989). *Priorities of educational testing and evaluation: The testimony of the CRESST national faculty* (CSE Tech. Rep. No. 304). Los Angeles: University of California, Center for the Study of Evaluation.

Heubert, J. & Hauser, M. (1999). *High stakes: Testing for tracking, promotion, and graduation*. Washington DC: National Academy Press.

Hoffman, J., Pennington, J., Assaf, L., & Paris, S. (1999). *High stakes testing in reading and its effects on teachers, teaching, and students: Today in texas, tomorrow?* Unpublished manuscript, University of Texas at Ausitn.

House, E. (1993). *Professional evaluation: Social impact and political consequences*. Newbury Park, CA: Sage.

Hoyle, R. & Smith, G. (1994). Formulating clinical research hypotheses as structural equation models: A conceptual overview. *Journal of Consulting and Clinical Psychology, Vol.62 (3)*, 429-440.

Hu, L. & Bentler, P. (1999). Cutoff criteria and fit indexes. *Structural Equation Modeling, 6(1)*, 1-55.

Igbaria, M., Gamers, T., & Davis, G. (1995). Testing the determinants of microcomputer usage via a structural equation model. *Journal of Management Information Systems, 11 (4)*, 87-114.

Keil, M. (1991). Managing MIS implementation: Identifying and removing barriers to use. Unpublished Doctoral Thesis, Harvard University.

Kirst, M. (1990). *Accountability: Implications for state and local policymakers*. Washington, DC: U.S. Government Printing Office.

Klein, S. P., Hamilton, L. S., McCaffrey, D. F., and Stecher, B. M. (2000). *What do test scores in Texas tell us?* (RAND Education Issue Paper) [Online]. Available: www.rand.org/publications/IP/IP202/

Kozol, J. (1991). *Savage inequalities: Children in America's schools*. New York: Crown.

Krathwohl, D. (1993). *Methods of educational and social science research*. White Plains, NY: Longman.

Leithwood, K. & Earl, L (2000). Educational accountability effects: An international perspective. *Peabody Journal of Education, 75 (4)*, 1-18.

Leithwood, K., & Steinbach, R. (1995). *Expert problem solving: Evidence from schools and district leaders.* Albany: State University of New York Press.

Lesnick, L., & Resnick, D.(1987). Standards, curriculum, and performance: A historical and comparative perspective. *Educational Researcher, 16 (9)*, 13-20.

Lessinger, L. (1970). *Every kid a winner: Accountability in education.* Palo Alto, CA: Science Research Associates.

Levin, H. (1974). A conceptual framework for accountability in education. *School Review, 82*, 363-391.

Lieberman, A. & Miller, L. (1999). *Teachers-transforming their world and their work*. New York, NY: Teachers College Press.

Linn, R. L. (2001). *The design and evaluation of educational assessment and accountability systems* (CSE Technical Report 539). Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing.

Linn, R., Baker, E., & Betebenner, D. (2002). Accountability systems: Implications of requirements of the No Child Left Behind Act of 2001. *Educational Researcher, 31 (6)*, 3-16.

Linn, R., Graue, M.., & Sanders, N. (1990). Comparing state and district results to national norms: The validity of the claims that "everyone is above average." *Educational Measurement: Issues and Practice, 9*(3), 5-14.

Lucas, H. & Spilter, V. (1999). Technology use and performance: A field study of broker workstations. *Decision Science, 30 (2)*, 291-311.

Macpherson, R. (1996). Educative accountability policy research: Methodology and epistemology. *Educational Administration Quarterly, 32 (1)*,

Madaus, G. (1988). The influence of testing on the curriculum. In L. Tanner (Ed.), *Critical issues in curriculum: 87th yearbook of the NSSE, Part 1* (pp. 83-121). Chicago: University of Chicago Press.

Malen, B., Ogawa, R., & Kranz, J. (1990). What do we know about school-based management? A case study of the literature-A call for research. In W.

Clune & J. Witte (Eds.), *Choice and Control in American Education* (Vol. 2, 289-342). New York: Falmer.

Mardia, K. (1970). Measures of multivariate skewness and kurtosis with application. *British Journal of Mathematical and Statistical Psychology, 28*, 205-214.

Marso, R. & Pigge, F. (1999). Educators' perceptions of the effectiveness of their schools' dissemination of standardized testing results. *American Secondary Education, 27(3),* 36-46.

Massell, D. (2001). The theory and practice of using data to build capacity: State and local strategies and their effects. In S. Fuhrman (Ed.), *From the capital to the classroom* (pp.148-169). University of Chicago Press.

McDonnell, L., & Elmore, R. (1987). Getting the job done: Alternative policy instruments. *Educational Evaluation and Policy Analysis, 9 (2)*, 133-152.

McGeoch, J. & Irion, A. (1952). *The psychology of human learning*. New York: Longmans.

McLaughlin, M., Shepard, L., & O'Day, J. (1995). *Improving education through standards-based reform.* Stanford, CA: National Academy of Education.

McMillan, J., & Schumacher, S. (1997). *Research in education: A conceptual introduction (4th ed.)*. New York: Addison-Wesley Educational Publishers, Inc.

McNeil, L. and Valenzuela, A. (2000). *The harmful impact of the TAAS system of testing in Texas: Beneath the accountability rhetoric.* Cambridge, MA: Harvard University Civil Rights Project.

Merriam, S. (1998). *Qualitative research and case study applications in education*. San Francisco, CA: Jossey-Bass.

Moore, G.C. & Benbasat, I. (1991). Development of an instrument to measure the perception of adopting an information technology innovation. *Information Systems Research 2(3), September*, 192-222.

Murnane, R. & Levy, F. (1996). *Teaching the new basic skills*. New York: The Free Press.

National Commission on Excellence in Education (1983). *A nation at risk: The imperative for educational reform.* Washington, DC: U.S. Government Printing Office.

Natriello, G. & Pallas, A. (1998). *The development and impact of high stakes testing.* Cambridge, MA: Harvard University Civil Rights Project.

Nickel, G. & Pinto, J. (1986). The computer attitude scale. *Computers in Human Behavior, 2*, 301-306.

Nuttall, D. (1994). Choosing indicators. In K. Riley & D. Nuttall (Eds.), *Measuring quality: Education indicators-United Kingdom and international perspectives* (pp. 17-40). Bristol, PA: Falmer.

O'Reilly, F. (1996). *Educational accountability: Current practices and theories in use.* Cambridge, MA: Harvard University, Consortium for Policy Research in Education.

Oakes, J. (1986). *Educational Indicators: A Guide for policymakers*. Santa Monica, CA: RAND Corporation, Center for Policy Research in Education.

Oakes, J. (1989). What educational indicators? The case for assessing the school context. *Educational Evaluation and Policy Analysis, 11 (2)*, 181-199.

Ogawa, R. & Collom, E. (2000). Using performance indicators to hold schools accountable: Implicit assumption and inherent tensions. *Peabody Journal of Education, 75 (4),* 200-215.

Oliver, C. (1991). Strategic responses to institutional processes. *Academy of Management Review, 16*, 145-179.

Olson, L. (1998). Study warns against reliance on testing data. *Education Week* (March 25):10.

Olson, L. (2000). Quality Counts 2001: Finding the right mix. *Education Week*, 1-14.

Orfield, R. (1999). Strengthening Title I: Designing a policy based on evidence. In G. Orfield & E. DeBray (Eds.), *Hard work for good schools: Facts not fads in Title I reform* (pp.1-20). Cambridge, MA: The Civil Rights Project, Harvard University.

Orlikowski, W. (1993). CASE tolls as organizational change: Investigating incremental and radical changes in systems development. *MIS Quarterly, 17(3)*, 309-340.

Patton, M. (1990). *Qualitative evaluation methods (2nd ed.)*. Thousand Oaks, CA: Sage.

Pipho, C. (1989). Accountability comes around again. *Phi Delta Kappan, 70 (9)*, 662-663.

Popham, W. (2000). Big change questions "Should large-scale assessment be used for accountability?" Answer: Depends on the assessment, silly! *Journal of Educational Change, 1*, 283-289.

Porter, A. (1994). National standards and school improvement in the 1990s: Issues and promise. *American Journal of Education, 102*, p.430.

Porter, A., & Smithson, J. (2001). *Defining, developing, and using curriculum indicators*. Research Report Series RR-048. Philadelphia, PA: Consortium for Policy Research in Education, University of Pennsylvania.

Porter, A., Smithson, J., & Osthoff, E. (1994). Standard setting for upgrading high school mathematics and science. In R. Elmore & S. Fuhrman (Eds.), *The governance of curriculum* (ASCD Yearbook)(pp. 138-166). Alexandria, VA: Association for Supervision and Curriculum Development.

Rallis, S. & MacMullen, M. (2000). Inquiry-minded schools: Opening doors for accountability. *Phi Delta Kappan, 81 (10)*, 766-773.

Richards, C. (1988). A typology of educational monitoring systems. *Educational Evaluation and Policy Analysis, 10 (2)*, 106-116.

Rivera, M. (1994). Accountability and educational reform in Rochester, New York. Doctoral dissertation, Harvard University.

Rogers, E.(1983). *Diffusion of Innovations, (3rd edition)*. New York: The Free Press.

Romberg, T.A., Zarinnia, A., & Williams, S. R. (1989). *The influence of mandated testing on mathematics instruction: Grade 8 teachers' perception (Monograph)*. Madison: University of Wisconsin, National Center for Research in Mathematical Sciences Education

Rosenberg, A. (1988). *Philosophy of social science*. Boulder, CO: Westview Press, Inc.

Rothman, R. (1995). *Measuring up: Standards, Assessment, and School Reform.* San Francisco: Jossey-Bass.

Rozum-Pratto, M., Gontarz, O., Flint, D., & Thomas, B. (1997). *Target teach*. Scottsdale, AZ: Evans Newton Inc.

Ruth, C. (2000). Applying a modified technology acceptance model to determine factors affecting behavior intentions to adopt electronic shopping on the world wide web: A structural equation modeling approach. Doctoral dissertation, Drexel University.

Scheerens, J. (1991). Process indicators of school functioning; A selection based on the research literature on school effectiveness. *Studies in Educational Evaluation, 17,* 371-403.

Schein, E. (1992). *Organizational culture and leadership*. San Francisco: Jossey-Bass.

Schwandt, T. (2000). Three epistemological stances for qualitative inquiry: interpretivism, hermeneutics, and social constructivism. In N. Denzin, & Y. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed., pp. 189-214). Thousand Oaks, CA: Sage Publications.

Selden, R. (1994). How indicators have been used in the USA. In K. Riley & D. Nuttall (Eds.), *Measuring quality: Education indicators-United Kingdom and international perspectives* (pp. 41-48). Bristol, PA: Falmer.

Shavelson, R. , McDonnell, L., & Oakes, J. (Eds.) (1989). *Indicators for monitoring mathematics and science education: A sourcebook* (Rep. No. R-3742-NSF/RC). Santa Monica, CA: RAND Corporation.

Shelor, L. (2000). A teacher appraisal invention process: Developing commitment to accountability. In B. Whitford & K. Jones (Eds.), *Accountability, assessment, and teacher commitment: Lesson from Kentucky's reform efforts* (pp.149-162). State University of New York Press.

Shepard, L. A. (1990). Inflated test score gains: Is it old norms or teaching the test? *Educational Measurement: Issues and Practices, 9(3)*, 15-22.

Skrla, L., Scheurich, J., & Johnson, J. (2001). Introduction: Toward a new consensus on high academic achievement for all children. *Education and Urban Society, 33*(3), 227-234.

Smith M., Edelsky, C., Draper, K., Rottenberg, C., & Cherland, M. (1989). *The role of testing in elementary schools.* Los Angeles: Center for Research on Educational Standards and Student Tests, UCLA.

Smith, M. (1988). Educational indicators. *Phi Delta Kappan, 69*, 487-491.

Smith, M. (1991). Put to the test: The effects of external testing on teachers. *Educational Researcher, 20 (5)*, 8-11.

Spring, J. (2000). *The American school 1642-2000*. Boston: McGraw-Hill

Stake, R. (1995). *The art of case study research*. Thousand Oaks, CA: Sage.

Stapleton, L. (October 10, 2001). Two-step structural equation modeling (Course materials of Structural Equation Modeling). University of Texas-Austin.

Steckler, A. (1989). The use of qualitative evaluation methods to test internal validity: An example of in a work site health promotion program. *Evaluation and the Health Professions, Vol. 12*, 115-133.

Steckler, A., McLeroy, K., Goodman, R., Bird, S., & McCormic, L. (1992). Toward integrating qualitative and quantitative methods: An introduction. *Health Education Quarterly, Vol. 19 (1)*, 1-8.

Stiggins, R. (1995). Assessment literacy for the 21st century. *Phi Delta Kappan, 77 (3)*, 238-245.

Stiggins, R. (1999). Assessment, student confidence, and school success. *Phi Delta Kappan, 81*, 191-198.

TAAS Manager (n.d.). The TAAS/TAKS challenge. Retrieved January 15, 2003, from http://www.taasmanager.com/challenge.cfm

Taylor, S., and Todd, P. (1995). Understanding IT usage: A test of competing models. *Information Systems Research, 6(2)*, 144-176.

Texas Education Agency. (n.d.). Student assessment division: TAKS information booklets. Retrieved January 15, 2003, from http://www.tea.state.tx.us/student.assessment/taks/booklets/index.html

Thorn, C. A. (2001). Knowledge management for educational information systems: What is the state of the field? *Education Policy Analysis Archives, 9 (47)*, as taken from http://olam.ed.asu/epaa/v9n47.

Urdan T. & Paris, S. (1994). Teacher' perception of standardized achievement tests. *Educational Policy, 8 (2)*, 137-156.

Venkatesh, V. & Davis, F. (2000). A theoretical extension of the Technology Acceptance Model: Four longitudinal field studies. *Management Science, 46(2),* 186-204.

Wagner, R. (1989). *Accountability in education: A philosophical inquiry*. New York: Routledge.

Wheelock, A. (2000). A new look at school accountability. In B. Whitford & K. Jones (Eds.), *Accountability, assessment, and teacher commitment: Lesson from Kentucky's reform efforts* (pp.179-198). State University of New York Press.

Wilson, S. & Floden, R. (2001). Hedging bets: Standards-based reform in classroom. In S. Fuhrman (ed.), *From the Capital to the Classroom* (pp.193-216). Univ. of Chicago Press.

Wynekoop, J. (1992). Strategies for implementation research: Combining research methods. In J.I. DeGross, J.D. Becker, & J.J. Elam (Eds.). *Proceedings of the 13th International Conference on Information Systems*. Dallas, TX.

Yin, R. (1994). *Case study research: Design and methods*. Thousand Oaks, CA: Sage.

Zmud, R. W. (1979). Individual differences and MIS success: A review of the empirical literature. *Management Science, 25(10)*, 966-979.

VITA

Sung-Kwan Yang was born in Goesan, Korea on January 10, 1969, the son of Woo-Seok Yang and Hak-Pyo Kim. After graduating from Myunmok High School in Seoul in 1987, he entered Korea University where he received the degree of Bachelor of Arts in February 1991. After receiving the degree of Master of Arts from Korea University in February 1993, he worked as a professional researcher (military service) at Korean Cultural Research Center for 5 years. In September 1998, he entered the Graduate School of University of Texas at Austin.

Permanent Address: 1343-3, Myunmok 4-dong, Jungrang-gu, Seoul, Korea, 131-204

This dissertation was typed by the author.